Anomalous Lending and Bank Risk-Taking

Abstract

Applying machine learning to over 150 loan attributes, we develop a novel anomaly score measure for syndicated bank lending that robustly predicts future bank risks. The loan-level anomaly score predicts higher borrower idiosyncratic risk and default probability, a greater likelihood of loans being sold at a discount, and wider bid-ask spreads in the secondary loan market. At the bank level, the loan-size-weighted average of anomaly scores positively predicts future bank risks for a period exceeding three years. The predictive power of the anomaly score is also stronger for banks with higher risk-taking incentives and those with greater complexity.

JEL classification: G17, G21

Keywords: machine learning; anomaly score; bank risks; syndicated lending; stress tests.

1. Introduction

Accurately estimating credit risk is of paramount importance in banking, often referred to as the "essence of banking" (Gorton and He, 2008). Syndicated loans, which serve as a cornerstone of corporate financing, have evolved into increasingly complex financial instruments. As evidence of this complexity, DealScan catalogs over 200 distinct loan attributes. Failing to grasp the intricacies of a bank's exposure in the syndicated loan market could result in significant economic and societal repercussions.[1] Despite extensive research into how bank loan contracts impact the behavior of borrowers and lenders,[2] the majority of empirical studies have focused on individual or limited subsets of loan attributes. Thus, the extant literature has not fully harnessed the high-dimensional nature of loan contracts, nor has it adequately explored the intricate interplay between various loan attributes.

In this study, we leverage machine learning technique to exploit the rich information set embedded in the vast array of syndicated bank loan attributes and develop a robust early-warning measure for identifying anomalous lending behavior. By navigating the complexity of high-dimensional syndicated loan data, we aim to enhance risk assessment practices in syndicated lending and contribute to the stability of the financial sector. Our approach not only addresses the pressing concerns of banking regulators but also provides valuable insights for market participants.

Given that identifying anomalous lending is conceptually equivalent to outlier detection, a well-studied topic in machine learning literature, we develop a real-time, loan-level

[1] For example, see Bernanke (1983), Diamond and Dybvig (1983), Calomiris and Mason (2003), Ivashina and Scharfstein (2010a), and Jermann and Quadrini (2012).

[2] For example, see Gorton and Kahn (2000), Dahiya, Saunders, and Srinivasan (2003), Ioannidou and Ongena (2010), Ivashina and Scharfstein (2010b), and Chan, Chen, and Chen (2013), among others.

anomaly score measure generated by a machine learning algorithm known as the Isolation Forest (Liu, Ting, and Zhou, 2008). This unsupervised algorithm builds on the simple intuition that outliers are easier to isolate and separate from normal observations that cluster together. It requires no predefined rules of normality nor a labeled training dataset, thereby circumventing the inherent difficulty of profiling normal loans to identify abnormal ones. Furthermore, the algorithm yields a continuous anomaly score ranging from 0 to 1 for each loan observation, representing the expected difficulty of isolating an individual loan from other loans that were originated. This anomaly score is not based on any distance measure between observations in the high-dimensional space. Thus, it does not suffer from the "curse of dimensionality" problem, where most observations have almost identical distances in high-dimensional spaces (e.g., Zimek, Schubert, and Kriegel (2012)). The algorithm is easily scalable for various applications involving large datasets.[3] Most recently, Isolation Forest has garnered research interest in both statistics and machine learning theory, as well as in its applications in medicine.[4]

We generate the anomaly score (hereafter referred to as the A-score) for each DealScan loan at origination by applying the Isolation Forest algorithm to quarterly samples of loans issued by all DealScan lenders in the past 12 months, encompassing 152 loan attributes.[5] We utilize the A-score at the origination of each loan as our anomalous lending measure.

[3]For example, LinkedIn employs the Isolation Forest algorithm in their Anti-Abuse AI Team to detect and prevent various types of abuse. See `https://engineering.linkedin.com/blog/2019/isolation-forest`.

[4]For example, Durairaj, Waterhouse, Mets, et al. (2023) successfully apply Isolation Forest to assign scores to proteins to predict structural outliers. Bates, Candès, Lei, Romano, and Sesia (2023) provide a rigorous theory of the false positive rate bound for outlier detection algorithms, including Isolation Forest, from the perspective of multiple testing.

[5]DealScan's new consolidated dataset contains lender-loan observations with 203 attributes, from which we manually select 152. Although we aim to include as many loan attributes as possible, we exclude the attributes that are perfectly correlated with others (e.g., firm name and firm identifier), provide no additional information (e.g., concatenation of other variables), or have almost entirely missing values. Online Appendix OA.2 details our A-score estimation process and variable selection.

We employ the quarterly decile rank of the A-score as the primary variable of interest in our empirical analyses to account for potential measurement errors in the A-score measure. Using ranking is advantageous, especially for skewed and non-normal data as observed in the raw A-scores (Lehmann and D'Abrera, 1975). Our results are also robust when using raw A-score values.[6]

We hypothesize that a bank's participation in a syndicated loan that significantly differs from other loans recently issued in the universe of syndicated loans may reflect the bank's risk-taking behavior. Therefore, the loan-level A-score, derived from applying the outlier detection machine learning algorithm to high-dimensional loan attributes of the universe of recently issued loans, may provide valuable insights into the credit quality of the loan.

To validate this hypothesis, we begin by documenting evidence showing that the A-score of a loan at origination, which quantifies how different the loan is (in a high-dimensional loan attribute space) compared with the universe of loans recently issued by all DealScan lenders, offers significant information about its borrower's current and future idiosyncratic risk and default probability. Furthermore, a loan-level analysis using a sample of DealScan loans from 2002 to 2018, matched with available secondary loan market trading data, reveals that loans with higher A-scores are more likely to be traded at a discount on the initial trading day. Moreover, the A-scores of loans sold in the secondary loan market exhibit a positive and significant correlation with larger discount sizes and wider bid-ask spreads (which serve as proxies for lower loan quality and higher loan risks). These findings remain robust when controlling for a battery of borrower characteristics, loan characteristics, year-quarter fixed

[6]Our robustness tests show that the main findings remain qualitatively unchanged when we use raw A-scores as the primary variable of interest.

effects, borrower industry fixed effects, loan purpose fixed effects, and loan type fixed effects. Therefore, in line with our hypothesis, the empirical evidence strongly suggests that loan-level A-scores indeed provide valuable information about borrowers and loan quality.[7]

In addition, using a bank-quarter panel dataset from 1997 to 2020, we investigate whether bank-level A-scores, calculated as the loan-amount-weighted average anomaly score at loan origination using all loans originated in the past 12 months that a given bank participates in, can predict future bank credit risks. We find that the quarterly decile rank of bank-level A-scores significantly and positively predicts the bank's future loan loss provisions for multi-year horizons ahead. The results remain robust when we control for various bank characteristics, year-quarter fixed effects, and even bank fixed effects.

Importantly, we observe that the predictive power of bank-level A-scores on future bank loan loss provisions increases in both magnitude and statistical significance as the predictive horizon approaches the value-weighted average maturity of the bank's loan portfolio, reaching its peak at approximately 12 quarters (or three years) ahead. A shift in the A-score rank from 1 to 10, on average, results in a 0.135-percentage-point increase in future loan loss provisions (relative to total assets) over the span of three years, equivalent to an additional expected loan loss of over $172 million annually. However, the predictive strength of bank-level A-scores diminishes and eventually dissipates as the predictive horizon surpasses the average maturity of the bank's loan portfolio. These findings suggest that bank-level A-scores can reliably predict future bank-level loan losses, particularly when the predictive horizon

[7] Berg, Saunders, and Steffen (2021) show that the volumes of both term loans and revolvers have increased over our sample period (see Figure 4 in their paper). We further interact a term loan indicator with loan-level A-score to investigate if there is any difference in the ability of the A-score to predict future borrower default risk for term loans. The results suggest no difference in the risk predictive power of the A-score between term loans and other loan types, such as credit lines.

aligns with loan maturity. If an anomalous h-year loan indicated by a high A-score rank at origination (time t) is more likely to become nonperforming later, we can reasonably expect this nonperforming risk to manifest around h years from the loan origination time t, which aligns with our findings.

In addition, we employ several approaches to investigate the nonlinearity, performance, and sources of the A-score's predictive power. First, to assess the nonlinearity of bank risk prediction, we demonstrate that the predictive power of anomalous lending for loan loss provisions primarily stems from higher A-score ranks rather than lower ranks. This finding aligns with our expectation that banks with exceptionally high A-scores contribute the most to the risk predictive power of the A-score. Regarding prediction performance, we show that the A-score provides incremental in-sample R^2 and reduces out-of-sample prediction errors, as measured by the root mean squared error (RMSE), compared to benchmark models that include only control variables and bank Z-scores. To investigate the source of predictive power, we classify all loan attributes into four categories: loan covenants, loan pricing, loan nonpricing, and borrower-related features. We then estimate four versions of the A-score using attributes from each category. Interestingly, we find that the A-score's primary risk predictive power originates from loan covenant attributes and marginally from loan pricing attributes, suggesting that the presence or absence of certain covenants may convey critical information regarding future bank loan losses.[8]

However, a loan's dissimilarity from other recently-issued loans in the universe of syndicated loans can also result from the lead bank's superior information about the borrower,

[8]However, it is worth noting that estimating the A-scores with a subset of loan attributes overlooks the potential complex interactions among attributes across different categories. We find that the interaction terms of the categorical A-scores also predict future bank risks.

acquired through due diligence and prior lending relationships. In such cases, the deviation in loan terms from the norm may not necessarily indicate risk-taking behavior and may not predict higher future loan losses or increased bank risks. On the contrary, when a bank participates in an anomalous syndicated loan as a lender, it typically lacks the same informational advantage as the lead arranger and may be more influenced by risk-taking incentives. Therefore, we hypothesize that the bank-level A-score, calculated based on loans in which a bank participates, will exhibit stronger predictive power for future bank risks compared to the bank-level A-score derived from loans lead-arranged by the bank.

To test this hypothesis, we decompose the bank-level A-score into two measures: one based on loans joined by the bank and the other on loans lead-arranged by the bank. Consistent with our hypothesis, we find that the predictive power of the bank-level A-score measure for future bank loan loss provisions primarily arises from loans in which the bank participated. In contrast, the A-score based on loans lead-arranged by the bank has little to no predictive power for future loan losses. These results confirm that lead-arranging anomalous loans does not predict adverse outcomes for the bank. It is the involvement in anomalous loans originated by other banks that represents a risk-taking activity of the bank, capable of predicting greater bank loan losses in the future.

Furthermore, we hypothesize that the predictive power of anomalous bank lending on future loan loss provisions may be influenced by bank risk-taking incentives and risk management practices. Specifically, banks that have recently experienced negative performance shocks may be compelled to take on more risks and engage in "abnormal" lending, consistent with the notion of banks gambling for resurrection. Similarly, anomalous lending might have a more pronounced impact on complex banks, where effective risk management is inherently

more challenging. Consistent with these hypotheses, we find that the A-score's predictive power for future bank loan losses is stronger for banks that have recently undergone poorer stock market performance and/or faced negative earnings surprises. Additionally, it is more pronounced for banks with greater complexity.

We conduct additional robustness tests on the predictive power of bank-level A-scores for other bank risk measures beyond loan loss provisions. We find that bank-level A-scores similarly predict greater future nonperforming loans (only nonperforming corporate loans, not nonperforming loans to individuals), higher future bank default probability, and higher future bank stock return idiosyncratic volatility. Therefore, our results affirm that bank anomalous lending in syndicated loans, as captured by bank-level A-score, is indeed associated with future bank loan losses and increased bank risks. We further find that the predictive power of bank A-score rank for future bank loan loss provisions concentrates on banks that rely more on interest income than noninterest income, alleviating the concern that banks are shifting to an "originate-to-distribute" model and not retaining syndicated loans. Furthermore, we find that the predictive power of bank-level A-score rank for future loan loss provisions is stronger during the Global Financial Crisis, but nonetheless remains positive and significant outside the crisis period. Splitting the sample period in half, we also find that the predictive power of the bank A-score rank remains stable in both periods.

Finally, we utilize the failures of banks in the Federal Reserve's forward-looking stress tests as a quasi-natural experiment to further validate the A-score as an early-warning tool for future bank risks. It is known that stress tests create incentives for banks to alter their future lending behaviors (Shapiro and Zeng, 2024). Through stacked-cohort difference-in-differences (DiD) and dynamic DiD estimations, we observe that banks failing the Federal

Reserve's Comprehensive Capital Analysis and Review (CCAR) stress tests significantly reduce their anomalous lending activities in syndicated loans, as measured by bank-level A-scores, in comparison to the control banks that passed the stress tests.[9] Given that banks failing the stress tests are mandated by the Federal Reserve to enhance their risk management practices, our findings affirm the A-score as a reliable early-warning bank risk measure. Furthermore, these results underscore the critical role of supervisory oversight in mitigating bank risk-taking activities, such as anomalous lending in syndicated loans.

Our paper contributes to the literature on syndicated lending and bank risks in several ways. The extant literature on syndicated lending has extensively explored various loan attributes related to loan outcomes and/or bank risks, including lead arranger's share (e.g., Leland and Pyle, 1977; Ivashina, 2009; Blickle, Fleckenstein, Hillenbrand, and Saunders, 2020), syndicate structure (e.g., Lin, Ma, Malatesta, and Xuan, 2012), bank-firm relationships (e.g., Ioannidou and Ongena, 2010; Bharath, Dahiya, Saunders, and Srinivasan, 2011) and type of lenders (e.g., Lim, Minton, and Weisbach, 2014), among others. To the best of our knowledge, this study is the first to adopt a holistic approach, considering over 150 loan attributes within a machine learning framework to generate an anomaly score for each syndicated loan in the DealScan universe.

We show that this score provides valuable insights into borrower and loan quality and exhibits strong predictive power for bank loan losses in the years ahead, particularly in proximity to the average maturity of the bank's loan portfolio. While impossible to establish causality, we contribute to the literature by demonstrating that a bank's anomalous lending

[9]Our finding that banks that fail stress tests reduce their anomalous lending activities in the syndicated loan market aligns with Gandhi and Purnanandam (2023), who show that following a stress test failure, a bank tends to become more similar to other banks in terms of equity returns, portfolio holdings, and exposure to macroeconomic factors.

activity in syndicated loans, as captured by our bank-level A-score, provides informative foresight into future bank risks.

Second, this study also contributes to the literature on bank risk-taking (e.g., Keeley, 1990; Saunders, Strock, and Travlos, 1990; Boyd and De Nicoló, 2005; Laeven and Levine, 2009). We show that while anomalous lending, on average, signifies bank risk-taking behavior, the primary predictive power for future bank risks stems from loans in which a bank participates as a lender rather than loans lead-arranged by the bank, where it has access to private borrower information due to its due diligence and past lending relationships (e.g., Allen, Gottesman, and Peng, 2012; Botsch and Vanasco, 2019).

Last but not least, our finding that failing the Federal Reserve's forward-looking stress tests reduces bank anomalous lending activities carries significant policy implications. Given that the A-score can serve as an early-warning tool for future bank risks, regulators should remain vigilant regarding banks' anomalous lending activities. This vigilance is crucial for assessing future bank loan performance and bank risks, ultimately contributing to the banking sector's resilience.

2. Hypothesis Development

Bank loan contracts contain high-dimensional information such as interest rate, maturity, collateral, covenants, loan purpose, syndication structure, and much more. The extant literature has extensively studied the impact of various exogenous shocks on bank loan contracting, as well as the effect of loan contracting on the borrower's and the lender's future performance and risks. It is known that banks strategically set loan contract terms in an-

ticipation of, and in response to, their borrowers' corporate actions and financial results. For example, Graham, Li, and Qiu (2008) find that, after corporate borrowers' financial restatements, banks charge higher loan spreads, allow for shorter loan maturities, require more covenant restrictions, and are more likely to demand collateral for the bank loans originating afterward. Chan et al. (2013) find that banks adjust loan contract terms when the borrower firm initiates clawback provisions by reducing interest rate and collateral requirements, and by increasing the financial covenants, performance pricing provisions, and loan maturities. Further, banks set loan contract terms to cater to the bank-firm relationship. For example, Ioannidou and Ongena (2010) shows that when firms switch banks, the new bank initially offers a lower interest rate, but this discount diminishes quickly. Prior lending relationships also affect loan rates and borrower's bankruptcy likelihood, and lead to larger shareholder value destruction for the lender should the borrower announce a default or bankruptcy (e.g., Dahiya et al., 2003; Bharath et al., 2011).

However, extant studies usually focus on a single aspect of the loan contract or, at best, a small subset of price and non-price loan attributes (e.g., loan spread, maturity, loan size, collateral usage, etc.), without taking a holistic approach to account for the high-dimensional and multi-faceted aspects of loan contracting and the possible complex interactions among various loan attributes. For example, loans to risky borrowers are typically characterized by higher loan spreads. However, they may also be contracted with normal loan spreads but stricter collateral requirements, tighter and/or more covenants on investments and financing, shorter maturity, and more frequent payment schedules. Given that the information contained in the high-dimensional loan contract space can be related to the lender's future loan losses and credit risks, and this association is neither linear nor straightforward,

we take a holistic approach using high-dimensional loan attributes and a machine learning methodology (Isolation Forest) to measure anomalous bank lending activities.

In our framework, measuring anomalous bank lending activities is conceptually equivalent to detecting outliers. This involves identifying loans that share fewer similarities in their loan attributes with other contemporaneous loans in the loan universe. These dissimilarities can span multiple attribute dimensions. For example, a bank may be lending to similar borrowers for the same loan purpose at the same time, while the loans differ from contemporaneous loans in the loan universe in terms of their sizes, spreads, collateral requirements, or covenant restrictions. In the context of syndicated loans, there may also be abnormally more or fewer banks involved in some loans compared to otherwise similar ones in the loan universe. An anomalous lending measure, based on how difficult it is to isolate a loan from other loans in a high-dimensional loan attribute space, can consolidate the high-dimensional attribute information to reflect how dissimilar the loan is relative to other contemporaneous loans made. Participating in loans that are very dissimilar to others may reflect bank risk-taking activity and could foreshadow future loan losses and bank risks.

Therefore, if syndicated loans participated in by a bank are substantially different from other contemporaneous loans in the syndicated loan market, there may be greater loan losses for this bank in the future. We develop the following hypothesis:

Hypothesis 1. *Anomalous lending activity predicts greater bank risks in the future.*

However, an 'anomalous' loan may not necessarily be of worse quality and lead to losses because deviations from the crowd can also be driven by private information of the lender; such informed 'abnormal' deviations cannot be reliably discerned from uninformed ones

beforehand. In this case, anomalous lending driven by private information will not be associated with higher future loan losses at the bank level. This follows the same rationale as the empirical finding that bold earnings forecasts from equity analysts are more accurate than herding forecasts due to private information (Clement and Tse, 2005). Therefore, it remains an empirical question whether anomalous lending could indicate greater or lower loan losses and future bank risks.

In a similar vein, a bank in the lead arranger role is expected to have superior information about the borrower due to the due diligence performed and the existing lending relationship (e.g. Allen et al., 2012; Balasubramanyan, Berger, and Koepke, 2019). Thus, anomalous loans may not pose a risk and potential loss to the lead arranger as the deviation can be due to the lead arranger's private information. In contrast, if an uninformed bank joins syndicated loans arranged by other banks, such activity is more likely to reflect risk-taking by the bank. We, therefore, expect that anomalous bank lending will have differential predictive power for the bank's future loan losses, depending on whether the loans are lead-arranged by the bank or joined by the bank. Specifically, we expect that participation in anomalous loans is more likely to be driven by bank risk-taking motives and thus has a stronger predictive power for future bank risks than lead-arranging anomalous loans. We develop the second hypothesis as follows:

Hypothesis 2. *The positive predictive power of bank anomalous lending on future bank risks mainly derives from the loans joined by a bank rather than from the loans lead-arranged by the bank.*

Moreover, we expect the predictive power of bank anomalous lending for future bank

risks to differ depending on the bank's recent performance and structural complexity. Banks that have experienced recent negative performance shocks are more likely to "gamble for resurrection" and take on more risks. Hence, anomalous lending by such banks is more likely to reflect risk taking, which can lead to greater loan losses in the future. In addition, anomalous lending can be more detrimental for complex banks because effective risk management that helps to reduce loan losses from risk taking is more difficult to achieve due to the complexity of these banks. We develop the third hypothesis as follows:

Hypothesis 3. *Anomalous lending activities are more informative about future bank risks for banks with poorer recent performance and banks that are more complex.*

Lastly, to further validate bank anomalous lending as an early-warning bank risk measure and to study the impact of supervisory oversight on bank anomalous lending, we use banks failing the Federal Reserve's forward-looking Comprehensive Capital Analysis and Review (CCAR) stress tests as quasi-natural experiments. We conduct stacked-cohort difference-in-differences (DiD) analyses to examine the changes in anomalous lending activities for failure banks relative to control banks after the stress-test-failure shocks.[10] Building on the insights from Shapiro and Zeng (2024), we conjecture that banks that failed a stress test should subsequently exhibit a significant reduction in anomalous lending activities relative to control banks.

[10]Failing a CCAR stress test leads to significant reputational damage and significant constraints on a bank's capital distribution plan. Failure banks are required to improve risk management, raise new equity capital and/or change their distribution plans.

3. Anomalous Lending and Outlier Detection

Measuring anomalous lending is conceptually equivalent to outlier detection, which is a well-studied topic in the machine learning literature. We apply an outlier detection algorithm, Isolation Forest, proposed by Liu et al. (2008) to isolate anomalous bank loans originating in the syndicated bank lending market. The algorithm is based on the simple intuition that outliers (anomalies) can always be isolated from the population easier than normal observations since anomalies are fewer than normal observations by definition.

To put the algorithm in a nutshell, we can imagine a two-dimensional space populated by some data points, each representing an observation. If we *randomly* cut the space into two along an axis (dimension) and repeat this process, an outlier far away from all others should be isolated (i.e., being the only data point in the subspace) faster than any inlier clustered with similar inliers. Each experiment of random cuts effectively grows (trains) an Isolation Tree, which fundamentally resembles a Binary Search Tree (BST). As we repeat the experiment of random cuts, the number of cuts needed to isolate a certain observation will quickly converge due to the properties of BST. Therefore, the Isolation Forest, a collection of Isolation Trees, can produce estimates of the average difficulty of separating any observation from the population, which can be used to derive an anomaly score measure.

We now formalize the estimation process of the Isolation Forest algorithm. Given an input dataset X of N observations and attribute set P, an Isolation Tree is constructed by recursively partitioning the dataset until all observations are isolated or a predetermined tree height l is reached. Specifically, we start with a random sample of X of size n ($n \leq N$). In each iteration, we first select a random attribute p from the attribute set P upon which

to split the sample, and then randomly select a split value q from the uniform distribution from the minimum to the maximum values of attribute p in the sample. The sample is then partitioned based on whether an observation's value of attribute p is larger than the split value q. In the next iteration, the same process is applied to the partitioned subsamples unless the subsample has only one observation (successful isolation) or the maximum number of iterations (tree height) is reached. If the iteration stops, the partitioned subsamples become the terminating nodes of the Isolation Tree; otherwise, they are internal nodes. The maximum tree height l is set based on the sample size n as $l = \text{ceiling}(\log_2(n))$, which is approximately the average height of a BST given size n. We restrict the tree height to the average because we are more interested in the anomalies that take less-than-average iterations to be isolated. An Isolation Forest of size M is just a collection of M Isolation Trees.[11]

Figure 1 provides a simple illustration of the Isolation Forest algorithm with two Isolation Trees in the context of syndicated loans.

[Insert Figure 1 about here]

The *anomaly score* of observation is the normalized expected number of iterations (path length in the tree) to isolate it from the sample. Given an observation x from the sample X and an Isolation Forest, we can estimate the expected path length $E(h(x))$ by taking the average of the path length of isolating x, $h(x)$, across all Isolation Trees in the Isolation Forest. The path length of x in an Isolation Tree can be easily determined if x belongs to a terminating node with only one observation.[12] However, if x belongs to a terminating

[11] We implement the algorithm in C++ as a Python extension for computation speed.
[12] In this case, the path length of x is the path length of the terminating node it belongs to.

node with more than one observation (due to iteration termination at the maximum tree height), we estimate its path length as the node's path length plus $c(k)$, the average path length of a BST of size k, where k is the number of observations in the terminating node.[13] Further, since $c(n)$ is the average of $h(x)$ given sample size n, we can use it to normalize $h(x)$. Therefore, the anomaly score s of an observation x is defined as:

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}}. \tag{1}$$

The anomaly score s is bounded by 0 and 1, and the higher the score, the fewer iterations on average it takes to isolate the observation, which implies a higher likelihood that the observation is an outlier.

The Isolation Forest provides at least three benefits for studying bank anomalous lending. First, the algorithm relies on neither predefined rules of normality nor labeled training datasets and produces a continuous measure of the observation's anomalousness compared to others in the sample. In the context of bank lending, it circumvents the inherently difficult profiling of "normal" loans in order to identify the "abnormal" ones. Second, the Isolation Forest algorithm, unlike traditional outlier detection algorithms, requires no parametrically defined distance between observations. Therefore, it is less affected by the "curse of dimensionality", in which most observations have almost identical distances in high-dimensional spaces, making it impossible to determine the threshold distance to identify outliers (see,

[13] In this case, the terminating node's path length is the number of iterations attempted to isolate x. Since we stop the search because it has taken more-than-average iterations, which implies that x is unlikely to be an outlier, there are still k observations in the terminating node. Therefore, we need to add the expected number of iterations to isolate x from the k observations in this terminating node. $c(k)$, the average path length of a BST of size k, is determined as $c(k) = 2H(k-1) - \frac{2(k-1)}{k}$, where $H(i)$ is the harmonic number and can be estimated by $\ln(i) + 0.5772156649$ (Euler's constant). See Preiss (1999) as discussed in Liu et al. (2008).

e.g., Zimek et al. (2012) for a survey). Third, the Isolation Forest algorithm can be easily scalable and does not consume extensive computing power since each Isolation Tree can be separately trained. This allows for efficient estimation of anomaly scores in the presence of a large-scale loans dataset and a high-dimensional loan attribute space.

4. Sample and Variable Construction

4.1. Measuring anomalous lending

To construct the bank lending anomaly score, we apply the Isolation Forest algorithm on a rich sample of bank loans constructed from the DealScan database. Specifically, at each quarter-end, we construct a training sample consisting of all loans originated in the past 12 months, with a total of 152 loan attributes. Online Appendix OA.2 describes how we screen all DealScan variables and encode categorical and textual values to construct the training sample that captures as much information as possible from the DealScan database.

We then apply the Isolation Forest algorithm to compute the loan-quarter anomaly scores with the same default parameters as in Liu et al. (2008), which are a subsample size of 256 and a forest size of 1,000 trees.[14] Figure 2 plots the quarterly distribution of loan-level anomaly scores at loan origination from 1991 to 2020.

[Insert Figure 2 about here]

[14] Liu et al. (2008) find that a subsample size larger than 256 provides minimal improvement for datasets with even more than 500,000 observations. Because the largest number of loan facilities originated per quarter in the DealScan database is 4,979 in 2018Q2, a subsample size of 256 captures around 5% of the loans in each Isolation Tree. Therefore, 1,000 trees in the Isolation Forest provide enough information to perform anomaly detection, and we use the default subsample size and forest size in this study. Nevertheless, we find that our results are robust to alternative subsample sizes and forest sizes, as shown in Table OA.3.2 in our Online Appendix.

To aggregate the loan-quarter anomaly scores to the bank-quarter level, we calculate the loan-amount-weighted average anomaly score (at loan origination) using all the loans originated in the past 12 months that a given bank participates in:

$$Bank\ A\text{-}score_{i,t} = \frac{\sum_j^{N_{i,t}} Loan\ A\text{-}score_j \times Loan\ Amount_j}{\sum_j^{N_{i,t}} Loan\ Amount_j}, \tag{2}$$

where $Bank\ A\text{-}score_{i,t}$ is the anomaly score of bank i at quarter t, $N_{i,t}$ is the number of loans that bank i participates in over the past 12 months up to the end of quarter t, $Loan\ A\text{-}score_j$ is the anomaly score of loan j at its origination,[15] and $Loan\ Amount_j$ is the dollar amount of loan j.

4.2. Sample construction and descriptive statistics

Our cross-sectional analysis of the predictive power of bank-level A-score uses a sample of U.S. bank holding companies' (BHC, or bank hereafter) lending data and financial characteristics from 1997 to 2020 conditional on the availability of all bank control variables.[16] Specifically, we collect all syndicated loans by U.S. banks from the Refinitiv LPC DealScan database, bank characteristics from Form FR Y-9C (consolidated financial statements for bank holding companies), and stock market data from the Center for Research in Security Prices (CRSP). For bank characteristics, we consider bank size, tier 1 capital ratio, the size of the bank's loan portfolio, deposits ratio, return on assets (ROA), market-to-book ratio (MTB), the bank's reputation and lending specialization. We merge bank characteristics

[15] Note that the anomaly score of a loan j varies over time because the training dataset is re-constructed at the end of each quarter consisting of all the loans originated in past 12 months. We use only the anomaly score estimated at loan origination to capture anomalous lending activities.

[16] Tier 1 capital is available since 1997.

and stock market data using the CRSP-FRB link table provided by the Federal Reserve Bank of New York.[17] Data on syndicated loans from DealScan are matched to bank characteristics based on the lenders' parent company using hand-matched bank name concordance files aggregated at the BHC level.[18]

[Insert Table 1 about here]

Table 1 reports the summary statistics for the main variables used in our empirical analyses. Definitions of the variables and data sources are provided in Table OA.1.1 in the Online Appendix. We winsorize all continuous variables by year-quarter at the 1st and 99th percentiles. Table OA.1.2 in the Online Appendix lists the top (bottom) five banks with the highest (lowest) levels of anomalous lending based on their average A-score decile rank. CIT Group Inc (Webster Bank NA) has the highest (lowest) average A-score rank of 9.15 (2.95) in the sample.

5. Anomalous Loans and Future Risks

We start with investigating whether the computed A-score of a loan contains information about its borrower's future risks. Specifically, we regress the natural logarithm of borrower's idiosyncratic stock return volatility on the quarterly decile rank of loan A-score, controlling for borrower characteristics such as firm size, Tobin's Q, ROA, asset tangibility, leverage, past stock return, and a wide range of loan characteristics including loan amount, maturity,

[17] The link table is available at https://www.newyorkfed.org/research/banking_research/datasets.html.

[18] The hand-matching process is carried out by two groups of research assistants independently and then carefully cross-checked by the authors for matching quality and consistency.

all-in-drawn spread, seniority, whether the loan is secured and whether it has performance pricing, the number of lenders, and the number of covenants. We also include the fixed effects for the time, borrower industry, loan purpose, and loan type. As many loans have multiple lead arrangers, we do not include lead arranger fixed effects, but we additionally control for the average reputation, syndicated loan specialization, and industry specialization of the lead banks. Standard errors are clustered at the year-quarter level. Panel A of Table 2 shows that loan-level A-score decile rank is positively and significantly associated with the contemporaneous and future (up to 12 quarters ahead) borrower idiosyncratic stock return volatility at the 1% level. This result suggests that the A-score of a loan is informative about the borrower's firm-specific risks. Moreover, in Panel B, we examine the relation between loan-level A-score rank and borrower firm's credit risk as measured by their Merton (1974) default probability and similarly find that loan-level A-scores positively and significantly predict borrower default probability. Further, in Table OA.1.3 in the Online Appendix, we interact loan-level A-score rank with the term loan indicator and find that the coefficient estimates of the interaction term are mostly insignificant. The results suggest that there is no difference in the risk predictive power of the loan-level A-score between term loans and other loan types, such as credit lines.

[Insert Table 2 about here]

We next investigate whether loan-level A-scores are related to worse loan outcomes in the secondary market. Using a sample of loans with matched secondary market trading data from 2002 to 2018, we regress the indicator of whether the loan is traded at a discount on the initial day of trading on loan A-score rank, after controlling for a wide range of loan

characteristics as before. As shown in column (1) of Table 3, we find that the loan-level A-score rank is positively associated with the likelihood of the loan being sold at a discount in the secondary market and is statistically significant at the 1% level. Column (2) shows that such a positive association between A-score rank and the likelihood of there being a secondary market trading discount remains qualitatively unchanged if we exclude those "distressed loans" that are traded at below 90% of face value, suggesting that it is not driven by extreme values of discount. In columns (3) and (4), we find that the loan's A-score rank is also positively and significantly related to the size of the trading discount at the 10% level and the bid-ask spread at the 1% level.[19]

[Insert Table 3 about here]

To summarize, the results from Table 2 and Table 3 suggest that loan-level A-scores are informative about borrower firm-specific risks and default probabilities; loans with higher A-scores at origination are more likely to be traded at a discount when sold to the secondary market; such loans are also more likely to have a larger discount size and wider bid-ask spread in the secondary market. Combined, Tables 2 and 3 provide empirical evidence suggesting that A-score foreshadows borrower risk and loan quality and hence may be used to predict future loan losses at the bank level.[20]

[19] Because the vast majority of loans sold in the secondary market are term loans, we do not interact the term loan indicator with loan-level A-score rank to test for heterogeneity.

[20] A potential concern is the relatively small number of loans traded in the secondary market compared to the number of loans in DealScan. The empirical evidence of Blickle et al. (2020) shows that 25% (35%) of loan shares are sold after initiation (over the loan duration). However, Blickle et al. (2020) also find that loans sold in the secondary market are less likely to become nonaccrual afterward than those retained by the lead banks due to concerns regarding their reputation and pipeline risk. Our finding that the A-score predicts trading discount and higher bid-ask spread even for these supposedly "good" loans in the secondary market further supports the A-score being an informative predictor.

6. Bank Anomalous Lending and Loan Loss Provisions

6.1. Baseline results

We now turn to investigate the predictive power of the A-score at the bank level for future loan loss provisions, which represents the expected loan losses of a bank over a certain period. Specifically, we estimate the following h-quarter-ahead predictive regressions:

$$LLP_{i,t+h} = \beta \text{Rank A-score}_{i,t} + \gamma X_{i,t} + \theta_t + \varepsilon_{i,t+h}, \tag{3}$$

where $LLP_{i,t+h}$ measures the loan loss provisions for bank i at time $t+h$, Rank A-score$_{i,t}$ is the quarterly decile rank of the bank-level A-score of bank i at time t (Equation (2)), $X_{i,t}$ is a vector of bank-level control variables at time t, and θ_t is the year-quarter fixed effects. We do not control for bank fixed effects because our focus is on the cross-sectional risk predictive power of bank-level A-scores.[21] For bank-level control variables, we include the natural logarithm of the book value of total assets, tier 1 capital ratio, loan portfolio size, deposits ratio, return on assets, market-to-book equity ratio, as well as the bank's reputation and lending specializations in the syndicated loan market and the borrower firms' industry.[22] Standard errors are clustered at the bank level.

[Insert Table 4 about here]

Table 4 presents the baseline results. For clarity, we report the results of 10 predictive

[21] Nevertheless, we also estimate the risk predictive regressions with control for bank fixed effects and report the results in Table OA.1.4 in the Online Appendix. The results remain qualitatively unchanged.

[22] We also estimate the predictive regression model without including any bank-level control variable and find qualitatively similar results.

regression models starting from 2-quarter-ahead up to 20-quarter-ahead predictions with a step of 2 quarters (i.e., $h = 2, 4, \ldots, 18, 20$). This wide range of prediction horizons is chosen because the potential risk of an anomalous loan takes time to manifest. Moreover, this range spans the maturity distribution of most syndicated loans and thus allows us to identify the timing of the risk predictive power of bank-level A-scores. We find that the coefficient estimates of banks' A-score rank are positive across all prediction horizons, increasing from 0.001 at $h = 2$ to 0.015 at $h = 12$ and then decreasing to 0.004 at $h = 20$. In terms of economic magnitude, increasing the bank-level A-score rank from 1 to 10, on average, leads to 0.135-percentage-point (0.015×9) higher future loan loss provisions relative to total assets 12-quarters ahead, which is about 64% of the sample mean of 0.211. Moreover, given the sample average bank size of 17.277 (natural logarithm of total assets in thousands), it implies an incremental expected loan loss of more than \$172 million per year ($e^{17.277} \times 0.135\% \times 4$), which is economically significant.

The associated distribution of t-statistics exhibits a similar inverted-U shape as the distribution of coefficient estimates. In other words, we find that the predictive power of bank A-score is strongest in magnitude and statistical significance for bank loan loss provisions 12 quarters in the future, which gradually declines as the prediction horizon moves away from 12 quarters in both directions.

An explanation for this inverted-U shape of bank A-score's predictive power for future loan loss provisions at different prediction horizons falls on loan maturity. Specifically, as a loan approaches its maturity, the risk of nonaccrual increases. Therefore, if an anomalous loan measured by a high A-score rank at origination is more likely to become nonperforming as time goes by, we reasonably expect the nonaccrual risk to peak around the maturity date

of the loan. For example, for a 3-year loan, a high A-score at origination (time t) should have the strongest predictive power for loan nonperforming risk (proxied by loan loss provisions) in around three years (time $t + 12$ quarters).

[Insert Figure 3 about here]

To illustrate this point, Figure 3 plots the coefficient estimates of bank-level A-score rank in the h-quarter-ahead predictive regressions of bank loan loss provisions from Table 4, and the distribution of bank loan portfolio's dollar-weighted maturity (in quarters). It is evident that, as expected, the distribution of the risk predictive power of bank A-score rank closely matches the distribution of the maturity of the bank loan portfolio. This evidence suggests that banks' A-scores at loan origination reliably predict greater loan losses around loan maturity – it neither predicts greater loan losses in the near or remote future nor generates predictive power that is evenly or randomly distributed across different prediction horizons.

In further robustness checks, we confirm that the findings are robust to double clustering standard errors at both the year-quarter and bank levels, as shown in Table OA.1.5 in the Online Appendix. In Table OA.1.6 in the Online Appendix, we find that the results remain qualitatively unchanged when we further control for bank loan pricing, measured by the all-in-drawn spread (in percentage points) weighted by the dollar loan amount of the syndicated loans that the bank participated in over the past 12 months.[23] Table OA.3.1 in the Online Appendix further presents the results of estimating Equation (3) using the raw bank-level A-scores, which are qualitatively similar to the results based on the decile rank of bank-level A-scores.

[23] At the loan level, we find that loan spread is positively related to A-score, albeit statistically insignificant.

To summarize, the empirical evidence in this section shows that anomalous bank lending activity strongly and reliably predicts greater loan loss provisions in the future, lending support to Hypothesis 1.

6.2. Nonlinearity, performance, and source of the predictive power

Since the construct of the A-score indicates that anomalous observations have a high A-score rank, we expect the predictive power of bank-level A-score for future loan loss provisions to mainly come from higher A-score ranks rather than lower ranks. We test this nonlinearity conjecture by replacing the decile rank of bank-level A-score with nine indicators for each specific decile rank value (excluding the lowest rank) in Equation (3).

[Insert Table 5 about here]

Table 5 reports the results. Consistent with our expectation, we find that the coefficient estimates tend to be larger for higher decile ranks of A-score. For easier comparison, Figure 4 plots the coefficient estimates from Table 5 of selected decile ranks (rank=2, 4, 6, 8, 10) of A-score for all prediction horizons, where darker color indicates higher ranks. It is clear that the predictive power of bank anomalous lending for loan loss provisions generally comes from higher A-score ranks than from lower ranks. This finding strongly suggests that it is the banks with high A-scores that contribute to the risk predictive power of A-scores.

[Insert Figure 4 about here]

Next, we conduct a series of tests to examine the in-sample and out-of-sample prediction performance of A-scores. Specifically, we focus on the 3-year-ahead predictions and estimate

four different model specifications that include: 1) control variables only, 2) bank Z-score and control variables, 3) bank A-score and control variables, and 4) both A-score and Z-score plus control variables. Panel A of Table 6 shows that the models with A-score improves the in-sample R^2 from 0.579 to 0.587. Also, the inclusion of bank Z-score does not affect the positive predictive power of bank A-scores for future loan loss provisions. These results confirm the incremental risk predictive power of bank A-scores. In Panel B of Table 6, we examine the out-of-sample prediction performance of A-score. Specifically, we use a rolling 3-year window to estimate the model specified by Equation (3), with which we then produce four out-of-sample 3-year-ahead predictions of bank loan loss provisions using the realized values of the independent variables (in Equation (3)) in each of the next four quarters as model inputs, respectively. We then pool all prediction errors together to compute the root-mean-squared-error (RMSE). We find that the inclusion of A-score reduces RMSE from 0.424 (controls only) and 0.417 (Z-score and controls) to 0.411 (A-score and controls) and 0.404 (A-score, Z-score and controls).

In summary, A-score provides incremental in-sample R^2 and reduces out-of-sample prediction errors, further confirming that bank anomalous lending measured by bank A-score can be an effective risk predictor for future loan loss provisions.

[Insert Table 6 about here]

Furthermore, we investigate the source of A-score's risk predictive power through a decomposition analysis. Since our A-score measure is estimated using over 150 loan attributes in the DealScan universe, unboxing the machine learning algorithm to identify the contributing factors of A-score is no easy task. To understand the source of A-score's risk predictive

power, we categorize all loan attributes into four categories, the ones related to 1) loan covenants, 2) loan pricing, 3) loan nonpricing features, and 4) borrower characteristics. For each category, we then separately estimate loan-level A-scores using all attributes classified into that category. We then similarly aggregate the A-score from each attribute category to bank-level using loan amount as weights. Next, we include the decile ranks of all four categorical A-score measures into the risk prediction model specified by Equation (3), replacing the single bank A-score estimated using all loan attributes.

[Insert Table 7 about here]

Table 7 shows that the risk predictive power of bank A-scores comes primarily from loan covenant attributes and marginally from loan pricing attributes. This suggests that the presence or nonpresence of certain covenants may be an important information source in predicting future bank loan losses. However, it is worth noting that estimating the A-scores using only a subset of loan attributes overlooks the potential complex interactions among loan attributes across different categories. Such cross-categorical interactions can also be an important source of information for predicting future bank risks.[24]

6.3. Loans joined vs. loans lead-arranged

As discussed earlier in Section 2, anomalous loans may not always lead to potential losses for a lead-arranger bank as the deviation from the norm can be driven by the lead arranger's superior private information rather than risk taking. Although private information and risk taking can both lead to deviation from the norm, we conjecture that high A-scores

[24]Table OA.1.7 in the Online Appendix shows that, when we include all interaction terms among the four categorical A-scores, most of these interaction terms are statistically significant at various levels.

of loans lead-arranged by a bank (who typically possesses private information about such loans) are, on average, more likely to reflect private information than high A-scores of loans participated by the bank (who does not have information advantage about such loans). We thus reasonably predict that the risk predictive power of A-score should mostly derive from loans participated in by a bank rather than from loans lead-arranged by the bank (Hypothesis 2). We examine this conjecture by replacing the bank-level A-score rank in Equation (3) with two bank-level A-score ranks separately estimated based on loans lead-arranged by the bank and loans joined by the bank in the past 12 months. All other control variables and model specifications remain unchanged. Table 8 presents the results.

[Insert Table 8 about here]

Consistent with our expectation, we find that the coefficient estimates of bank A-score rank based on loans joined share similar size and statistical significance as in the baseline results in Table 4. In sharp contrast, the coefficient estimates of the bank A-score rank based on loans lead-arranged are statistically insignificant across all prediction horizons. They are even negative (indicating that greater A-score rank may be associated with lower future loan losses) most of the time. These results suggest that the relationship between bank anomalous lending activities and future bank risks hinges crucially on whether such activities are driven by risk-taking or private information. Deviation from the norm, if driven by superior information, does not necessitate a bad outcome, but participation in anomalous loans initiated by others is risky. The findings lend strong support to Hypothesis 2.

6.4. Performance shocks, bank complexity, and the risk predictive power of bank A-score

As discussed in Section 2, we also conjecture that the risk predictive power of anomalous lending may be moderated by the risk-taking incentives and risk management practices of a bank. Specifically, we expect the predictive power of bank A-score for future loan loss provisions to be stronger when the bank has experienced negative performance shocks, which puts the bank under increasing pressure to take on more risk and lend "abnormally" in a gamble for a quick turnaround. Furthermore, anomalous lending can be more detrimental for complex banks where effective risk management is more challenging to achieve.

6.4.1. Performance shocks

To examine whether bank anomalous lending induced by recent poor performance is riskier, we consider two bank performance measures based on the stock market and accounting metrics. The market-based measure is the past-year buy-and-hold abnormal return computed using the market model. For the accounting measure, we use the earnings surprise computed using the difference between actual earnings and the average analyst's forecast, normalized by the stock price. We then include the bank's performance measure and its interaction with the bank's A-score rank in the baseline model as in Equation (3). Table 9 reports the results.

[Insert Table 9 about here]

We find that the coefficient estimates of bank A-score rank remain positive and statistically significant as in the baseline results. More importantly, consistent with Hypothesis 3,

we find that the coefficient estimates on the interaction terms between A-score rank and two bank performance measures are negative and statistically significant around the average maturity of all banks' loan portfolios. These results suggest that the anomalous lending activities of banks that experienced poorer recent performance are associated with even higher loan loss provisions in the future, especially when the anomalous loans approach their maturity. This relationship is present regardless of whether the recent bank performance is measured using market or accounting data.

6.4.2. Bank complexity

Next, we test whether the predictive power of anomalous lending for future loan loss provisions is stronger for more complex banks. We capture bank complexity using a simple measure based on the number of non-missing items reported in FR Y-9C. Complex banks arguably have more items to report, and thus, the ratio of the number of non-missing items to the number of total items in FR Y-9C provides an intuitive and parsimonious measure of a bank's complexity. This measure is similar to the measure proposed by Chen, Miao, and Shevlin (2015), the ratio of the number of non-missing items to the number of total items in Compustat, which is used to capture the "disclosure fineness" of industrial firms as it reflects the level of disaggregation of accounting data in annual reports. In the context of banking, this ratio captures more of the complexity of a bank's business. As of February 2022, Compustat has 948 distinct line items for firms' annual reports, while FR Y-9C has over 2,300 line items. Apart from the standard balance-sheet items, banks are required to report in great detail about their off-balance-sheet activities, including, for example, unused loan commitments, standby letters of credit, revolving underwriting facilities, and credit

derivatives, among other things. A missing item naturally implies that the reporting bank needs to have such business. Hence, the more non-missing items reported by a bank, the more complex the bank's business activities typically are.

Given that the number of non-missing items reported in FR Y-9C may be positively correlated with bank size, we perform an orthogonalization of bank complexity to bank size by regressing the complexity ratio on bank size. We then use the residual from this regression as our bank complexity measure to be included additionally in Equation (3) (as well as its interaction term with bank A-score rank). Table 10 presents the results.

[Insert Table 10 about here]

We find that the coefficient estimates of bank A-score again remain positive and statistically significant, similar to the baseline results. Importantly, consistent with Hypothesis 3, the interaction terms between A-score rank and bank complexity are positive and statistically significant, particularly around the average maturity of bank loan portfolio. These results indicate that the anomalous lending activities of complex banks are associated with even higher loan loss provisions in the future, more so when the anomalous loans approach their maturity.

To summarize, we find that bank anomalous lending activities are more informative about future bank risks for banks with poorer recent performance and banks that are more complex, lending empirical support to Hypothesis 3.

6.5. Additional robustness results

For robustness concern, we further validate the predictive power of bank A-score for other bank risk measures beyond bank loan loss provisions. These alternative bank risk measures include actual nonperforming loans, default probability, and idiosyncratic stock return volatility; the latter two measures are market-based measures. Table 11 shows that a higher bank-level A-score rank similarly predicts greater actual nonperforming loans, higher bank default probability (measured using both the original Merton (1974) model via a KMV iterative approach, and the Nagel and Purnanandam (2020) modified model for banks), and larger idiosyncratic stock return volatility. Moreover, the size and statistical significance of the coefficient estimates of bank A-score rank exhibit a similar inverted-U shape as in Table 4, peaking at around the bank loan portfolio's average maturity (about 12 quarters). These results, based on employing various market-based and non-market-based bank risk measures, suggest that bank A-score can indeed reliably and robustly predict future bank risks.

[Insert Table 11 about here]

Moreover, in Table OA.1.8 in the Online Appendix, we decompose actual nonperforming loans into loans extended to firms and loans extended to individuals. We find that the predictive power of bank A-score rank for actual nonperforming loans comes exclusively from corporate loans, which is consistent with the fact that our estimation of A-score is conducted on corporate syndicated loans. Additionally, given the concern that some banks may shift to an "originate-to-distribute" model and retain less syndicated loans, we conjecture that the predictive power of bank A-score rank for future bank loan loss provisions should be higher

for banks with more traditional lending businesses. In Table OA.1.9 in the Online Appendix, we test this conjecture based on the composition of the bank's income. We find that the predictive power of bank A-score rank comes mainly from banks with above year-quarter median ratio of interest to noninterest income. This result aligns with our expectation that bank anomalous lending only predicts greater future bank loan loss provisions when the bank's income relies more on the lending business than other activities.

Lastly, we examine whether the predictive ability of bank A-score holds both inside and outside of the Global Financial Crisis (GFC) period. To achieve this purpose, we interact with the GFC indicator (which takes the value of one for the year-quarters from 2007Q1 to 2009Q4 and equals zero otherwise) and the bank-level A-score rank. Table OA.1.10 in the Online Appendix shows that the predictive power of bank-level A-score rank for future loan loss provisions is stronger during the GFC period. Nonetheless, the coefficient estimates of bank-level A-score rank remain positive and significant, suggesting significant predictive power beyond the GFC period. Moreover, we split the sample period into two equal-length sub-periods and create an indicator variable PostMid which takes the value of one for the observations in the more recent sub-period and equals zero otherwise. Table OA.1.11 in the Online Appendix shows that the coefficient estimates of the interaction term are statistically insignificant while the bank-level A-score rank remains positive and significant, which implies that the predictive power of bank-level A-score rank for future loan loss provisions is relatively stable over time.

7. Impact of Stress Test Failure on Bank Anomalous Lending Activities

In this section, we examine the impact of banks failing the Federal Reserve's supervisory stress tests on their (anomalous) lending activities. In the wake of the Global Financial Crisis, the supervisory bank stress tests, which started with the Supervisory Capital Assessment Program (SCAP) in 2009, were introduced by the Federal Reserve as a forward-looking supervisory tool to ensure that banks have enough capital to survive adverse economic shocks. Most of the tested bank holding companies failed the 2009 SCAP test. In 2011, the Federal Reserve began conducting the annual Comprehensive Capital Analysis and Review (CCAR) stress test to determine the capital adequacy of large bank holding companies. Banks that fail a CCAR stress test are required to improve risk management, raise new capital, and/or change their capital distribution plans. Thus, a stress test failure is a bad public signal and significantly damages the bank's reputation. An overview of the participating banks and the outcomes of the stress tests across different rounds are shown in Table OA.1.12 in the Online Appendix.

The earlier sections show that bank-level A-score serves as an early-warning predictor for future bank risks. Given the exogenous negative shock to a bank when it failed a supervisory stress test and the mandate to improve risk management practice, we expect the failed bank will take action to decrease its anomalous lending activities as measured by its A-score. We thus perform a stacked-cohort difference-in-differences (DiD) estimation using stress test failure as the treatment to investigate the effect of failing a supervisory stress test on bank-level A-score.

Specifically, we use only the CCAR tests from 2012 through 2019 in our stacked-cohort DiD analyses. This is because most of the tested banks failed in the SCAP test in 2009 and the result of the CCAR test in 2011 has not been publicly disclosed. Thus, our sample consists of all banks that have ever participated in a CCAR test from 2012 to 2019. We construct a subsample (cohort) for each round of the CCAR tests. We classify banks that failed (passed) the focal round of CCAR as the treated (control) banks. If a bank failed multiple rounds of CCAR, we only consider the first stress test failure of the bank and remove the bank from the subsamples of subsequent test rounds. Since a bank may have changed its lending behavior after a stress test failure, considering subsequent stress test failures of the bank can bias our estimation results. For each cohort, We use an event window surrounding each CCAR test round from two years before to two years after the year of the focal test round (excluding the focal test year). We then stack all the cohorts together and estimate a standard stacked-cohort DiD regression specification as follows:

$$\text{A-score}_{i,c,t} = \beta_1 Treat_{i,c} + \beta_2 Post_{c,t} + \beta_3 Treat_{i,c} \times Post_{c,t} + \lambda_i + \theta_t + \varepsilon_{i,c,t}, \tag{4}$$

where $\text{A-score}_{i,c,t}$ is the A-score of bank i in cohort c at year t, $Treat_{i,c}$ is an indicator variable that equals 1 if bank i in cohort c is a treatment bank and equals 0 otherwise, $Post_{c,t}$ is an indicator variable that equals 1 if year t in cohort c is after the focal test year and equals 0 otherwise, λ_i is the bank fixed effects, and θ_t is year fixed effects. We deliberately exclude bank control variables due to the concern of a "bad control" problem, where the treatment event (stress test failure) may affect the control variables (e.g., liquidity, loan growth, capital ratio, etc.) in the post-treatment period.

For robustness, we also estimate a similar regression specification where we replace bank fixed effects with the more conservative cohort-bank fixed effects and year-fixed effects with cohort-year fixed effects as follows:

$$\text{A-score}_{i,c,t} = \beta Treat_{i,c} \times Post_{c,t} + \omega_{i,c} + \gamma_{c,t} + \varepsilon_{i,c,t}, \tag{5}$$

where the $Treat_{i,c}$ and $Post_{c,t}$ indicators are absorbed by the cohort-bank fixed effects $\omega_{i,c}$ and cohort-year fixed effects $\gamma_{c,t}$, respectively. When estimating the above two DiD regression specifications, we cluster standard errors at the bank level or at both the bank and cohort levels. The results are reported in Panel A of Table 12. We find that treatment banks that failed a stress test, on average, reduced their anomalous lending activities (as measured by their bank A-score) relative to control banks that passed the stress test.

[Insert Table 12 about here]

Moreover, we employ a dynamic DiD regression framework to examine whether the treatment effect of stress test failures on bank-level A-score is driven by potential nonparallel lending trends between the treatment and control banks prior to the release of a stress test outcome. Specifically, we include the focal stress test year in the sample and use the following regression specification to identify the exact timing of the treatment effect:

$$\begin{aligned}\text{A-score}_{i,c,t} = {} & \beta_1 Treat_{i,c} \times d_{-1,c,t} + \beta_2 Treat_{i,c} \times d_{0,c,t} \\ & + \beta_3 Treat_{i,c} \times d_{1,c,t} + \beta_4 Treat_{i,c} \times d_{2,c,t} \\ & + \beta_5 d_{-1,c,t} + \beta_6 d_{0,c,t} + \beta_7 d_{1,c,t} + \beta_8 d_{2,c,t} \\ & + \beta_9 Treat_{i,c} + \lambda_i + \theta_t + \varepsilon_{i,c,t},\end{aligned} \tag{6}$$

where we replace the single $Post_{c,t}$ indicator in Equation 4 with year-specific indicators $d_{j,c,t}$ ($j \in \{-1, 0, 1, 2\}$). Specifically, $d_{j,c,t}$ equals 1 if year t is the jth year relative to the event year in cohort c and equals 0 otherwise. Year -2 in each cohort is used as the reference year. For robustness, we also estimate the following dynamic DiD regression specification with the cohort-bank and cohort-year fixed effects:

$$\begin{aligned}\text{A-score}_{i,c,t} = {} & \beta_1 Treat_{i,c} \times d_{-1,c,t} + \beta_2 Treat_{i,c} \times d_{0,c,t} + \beta_3 Treat_{i,c} \times d_{1,c,t} \\ & + \beta_4 Treat_{i,c} \times d_{2,c,t} + \omega_{i,c} + \gamma_{c,t} + \varepsilon_{i,c,t}.\end{aligned} \tag{7}$$

This dynamic DiD regression framework enables us to identify the timing of the onset of the treatment effect. If the reduction in bank-level A-score is indeed caused by bank stress test failures, we should expect no difference-in-differences between the treatment and control banks prior to the release of the stress test outcome, which implies a statistically insignificant β_1. Any reduction in bank-level A-score due to stress test failures should be captured by negative and significant DiD estimates in and after the event year (i.e., β_2, β_3, and β_4).

Panel B of Table 12 reports the dynamic DiD estimation results. We find that across all regression specifications, the coefficient estimate of β_1 is statistically insignificant, which suggests that treatment and control banks have parallel A-score trends prior to the release of

stress test outcomes. Importantly, the negative and significant treatment effect of stress test failures on bank-level A-score is observed only on and after the release of stress test outcomes. Taken together, these findings suggest that stress test failures prompt failing banks to reduce their abnormal lending activities. Given that stress tests are a forward-looking supervisory tool and failure banks are mandated to improve their risk management, these findings not only validate the bank-level A-score measure as an early-warning measure for future bank risks but also highlight the importance of regulatory oversights on anomalous bank lending activities.

8. Conclusion

In this paper, we develop a new measure for anomalous lending in syndicated loans, A-score, that reliably predicts future bank loan loss provisions and other bank risk measures. We apply a machine learning framework to over 150 loan attributes to develop the A-score measure for each loan in the DealScan universe. We show that the A-score measure at the loan level is associated with higher borrower firm risks and worse loan outcomes such as a higher likelihood of being sold in the secondary loan sales market at a large discount and a wider bid-ask spread. At the bank level, the value-weighted average of loan-level A-scores positively predicts bank loan loss provisions for several years ahead, with the predictive power peaking at around the average maturity of bank loan portfolios (i.e., a prediction horizon of around three years ahead). The risk predictive power of bank-level A-score is nonlinear and comes mainly from higher A-scores. Moreover, the predictive power derives mainly from the loans that the bank joins rather than from the loans that it leverages and is

stronger for banks that have stronger risk-taking incentives due to poor recent performance and for more complex banks where effective risk management is harder to achieve. We further show that banks that failed the Federal Reserve's forward-looking CCAR stress tests exhibit a significant reduction in their anomalous lending activities as measured by their bank A-score, relative to the control banks that passed the stress tests.

Combined, these findings suggest that bank-level anomalous lending, as captured by the machine-learning-based A-score measure developed from high-dimensional loan attribute data, can serve as an informative, early-warning bank risk measure for banks, investors, and regulators. There are pertinent policy implications from our findings based on this new early-warning indicator as supervisory monitoring of bank anomalous lending activities can be particularly important to rein in future bank loan losses for a more resilient banking sector.

References

Allen, L., Gottesman, A. A., Peng, L., 2012. The impact of joint participation on liquidity in equity and syndicated bank loan markets. *Journal of Financial Intermediation* 21, 50–78.

Balasubramanyan, L., Berger, A. N., Koepke, M. M., 2019. How do lead banks use their private information about loan quality in the syndicated loan market? *Journal of Financial Stability* 43, 53–78.

Bates, S., Candès, E., Lei, L., Romano, Y., Sesia, M., 2023. Testing for outliers with conformal p-values. *The Annals of Statistics* 51, 149 – 178.

Berg, T., Saunders, A., Steffen, S., 2021. Trends in corporate borrowing. *Annual Review of Financial Economics* 13, 321–340.

Bernanke, B. S., 1983. Nonmonetary effects of the financial crisis in the propagation of the great depression. *The American Economic Review* 73, 257–276.

Bharath, S. T., Dahiya, S., Saunders, A., Srinivasan, A., 2011. Lending relationships and loan contract terms. *The Review of Financial Studies* .

Blickle, K., Fleckenstein, Q., Hillenbrand, S., Saunders, A., 2020. The myth of the lead arranger's share. Tech. Rep. 922, Federal Reserve Bank of New York.

Botsch, M., Vanasco, V., 2019. Learning by lending. *Journal of Financial Intermediation* 37, 1–14.

Boyd, J. H., De Nicoló, G., 2005. The theory of bank risk taking and competition revisited. *The Journal of Finance* 60, 1329–1343.

Calomiris, C. W., Mason, J. R., 2003. Consequences of bank distress during the Great Depression. *American Economic Review* 93, 937–947.

Chan, L. H., Chen, K. C. W., Chen, T.-Y., 2013. The effects of firm-initiated clawback provisions on bank loan contracting. *Journal of Financial Economics* 110, 659–679.

Chen, S., Miao, B., Shevlin, T., 2015. A new measure of disclosure quality: The level of disaggregation of accounting data in annual reports. *Journal of Accounting Research* 53, 1017–1054.

Clement, M. B., Tse, S. Y., 2005. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance* 60, 307–341.

Dahiya, S., Saunders, A., Srinivasan, A., 2003. Financial distress and bank lending relationships. *The Journal of Finance* 58, 375–399.

Diamond, D. W., Dybvig, P. H., 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91, 401–419.

Durairaj, J., Waterhouse, A. M., Mets, T., et al., 2023. Uncovering new families and folds in the natural protein universe. *Nature* 622, 646–653.

Gandhi, P., Purnanandam, A., 2023. United they fall: Bank risk after the financial crisis. *SSRN Electronic Journal* .

Gorton, G., Kahn, J., 2000. The design of bank loan contracts. *The Review of Financial Studies* 13, 331–364.

Gorton, G. B., He, P., 2008. Bank credit cycles. *Review of Economic Studies* 75, 1181–1214.

Graham, J. R., Li, S., Qiu, J., 2008. Corporate misreporting and bank loan contracting. *Journal of Financial Economics* 89, 44–61.

Ioannidou, V., Ongena, S., 2010. "Time for a change": Loan conditions and bank behavior when firms switch banks. *The Journal of Finance* 65, 1847–1877.

Ivashina, V., 2009. Asymmetric information effects on loan spreads. *Journal of Financial Economics* 92, 300–319.

Ivashina, V., Scharfstein, D., 2010a. Bank lending during the financial crisis of 2008. *Journal of Financial Economics* 97, 319–338.

Ivashina, V., Scharfstein, D., 2010b. Loan syndication and credit cycles. *The American Economic Review* 100, 57–61.

Jermann, U., Quadrini, V., 2012. Macroeconomic effects of financial shocks. *The American Economic Review* 102, 238–271.

Keeley, M. C., 1990. Deposit insurance, risk, and market power in banking. *The American Economic Review* 80, 1183–1200.

Laeven, L., Levine, R., 2009. Bank governance, regulation and risk taking. *Journal of Financial Economics* 93, 259–275.

Lehmann, E. L., D'Abrera, H. J., 1975. Nonparametrics: Statistical methods based on ranks. Holden-Day.

Leland, H. E., Pyle, D. H., 1977. Informational asymmetries, financial structure, and financial intermediation. *The Journal of Finance* 32, 371–387.

Lim, J., Minton, B. A., Weisbach, M. S., 2014. Syndicated loan spreads and the composition of the syndicate. *Journal of Financial Economics* 111, 45–69.

Lin, C., Ma, Y., Malatesta, P., Xuan, Y., 2012. Corporate ownership structure and bank loan syndicate structure. *Journal of Financial Economics* 104, 1–22.

Liu, F. T., Ting, K. M., Zhou, Z., 2008. Isolation forest. In: *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422.

Merton, R. C., 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* 29, 449–470.

Nagel, S., Purnanandam, A., 2020. Banks' risk dynamics and distance to default. *The Review of Financial Studies* 33, 2421–2467.

Preiss, B. R., 1999. Data Structures and Algorithms. John Wiley & Sons, Inc., USA, first ed.

Saunders, A., Strock, E., Travlos, N. G., 1990. Ownership structure, deregulation, and bank risk taking. *Journal of Finance* 45, 643–654.

Shapiro, J., Zeng, J., 2024. Stress testing and bank lending. *The Review of Financial Studies* 37, 1265–1314.

Zimek, A., Schubert, E., Kriegel, H.-P., 2012. A survey on unsupervised outlier detection in high-dimensional numerical data. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 5, 363–387.

Fig. 1. **Example Illustration of Isolation Forest**

Figure 1 illustrates an Isolation Forest of two Isolation Trees. Each Isolation Tree is of size n such that the maximum height is 3. At each iteration, a random loan attribute p is selected as well as a random split value q from the uniform distribution from the minimum to the maximum values of p in the subsample, with which the sample is splitted into two. Terminating nodes with successful (unsuccessful) isolation are coloured blue (red). In this example, loan j has a path length of 3 in the first Isolation Tree and a path length of 2 in the second. Therefore, its average path length from this forest is estimated as $(3 + 2)/2 = 2.5$.

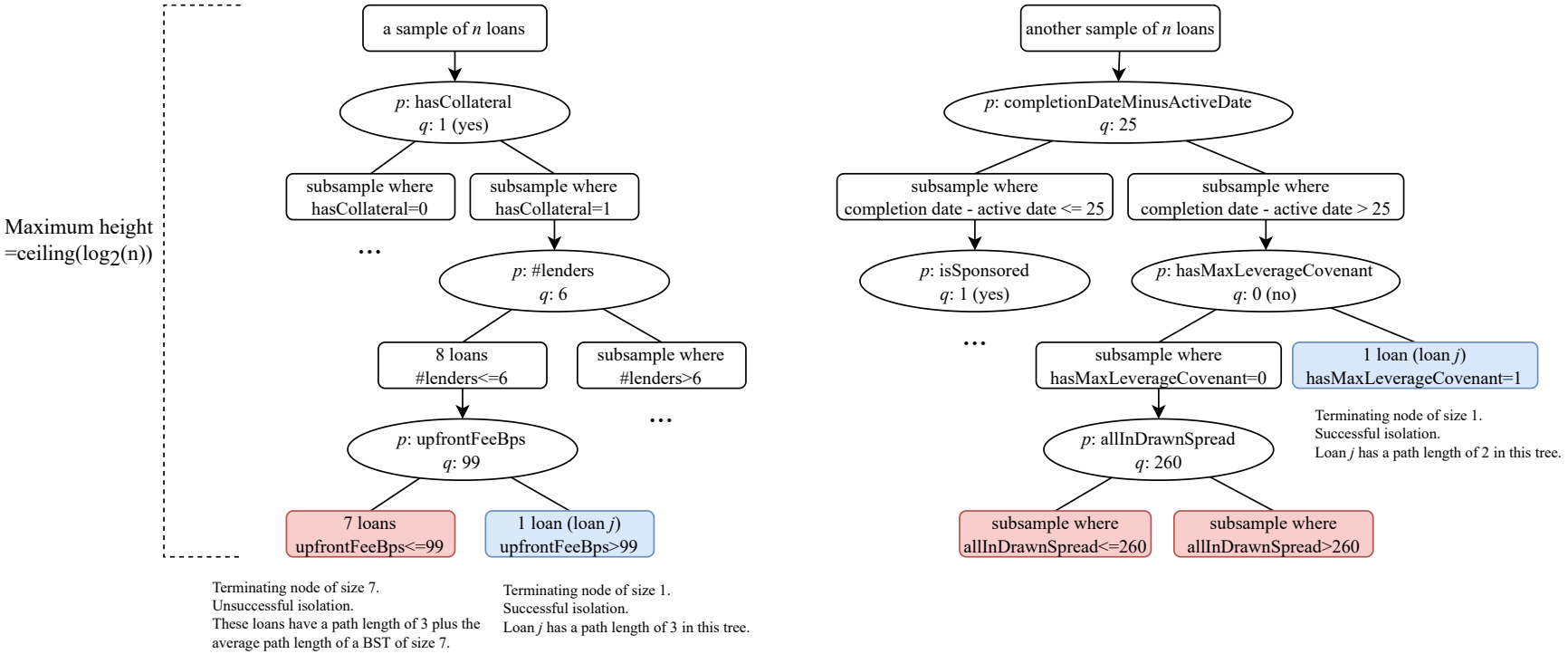


Fig. 2. **Distribution of A-score 1991-2020**

Figure 2 plots the quarterly distributions of loan-level A-scores from 1991 to 2020. The loan-level A-scores are estimated using a quarterly rolling window with a length of 12 months. The plot shows the quarterly distribution of loan A-scores at loan origination.

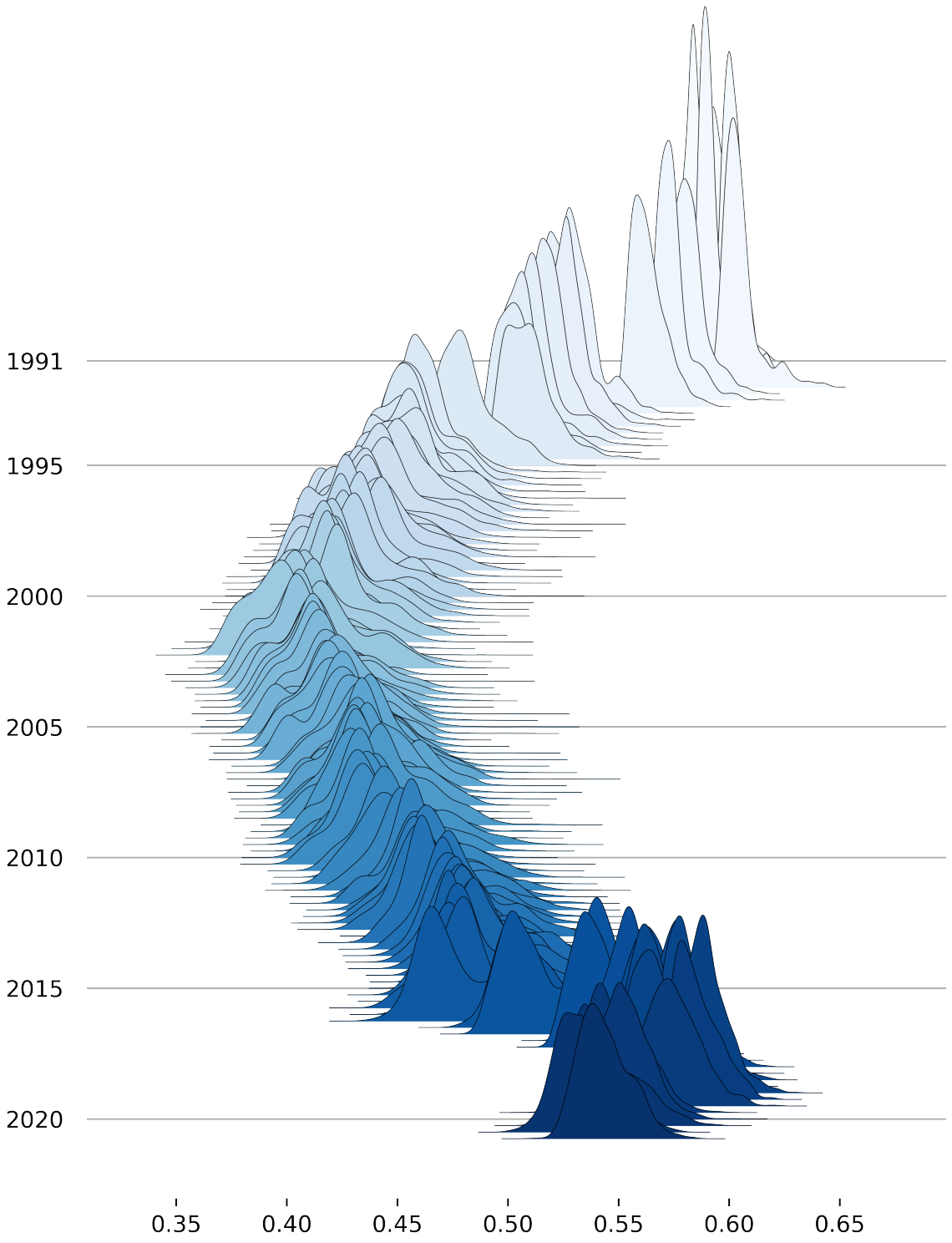


Fig. 3. **Predictive Power of Bank-level A-score and Loan Maturity**

Figure 3 plots the coefficient estimates of bank-level A-score rank in the h-quarter-ahead prediction of future bank loan loss provisions from Table 4 and the distribution of bank loan portfolio's dollar-weighted maturity (in quarters, capped at 20 in the plot). We show that bank anomalous lending captured by A-score exhibits increasing predictive power as the forecast horizon h (x-axis) is closer to the average maturity of the bank's loan portfolio, peaked at around 12 quarters ahead.

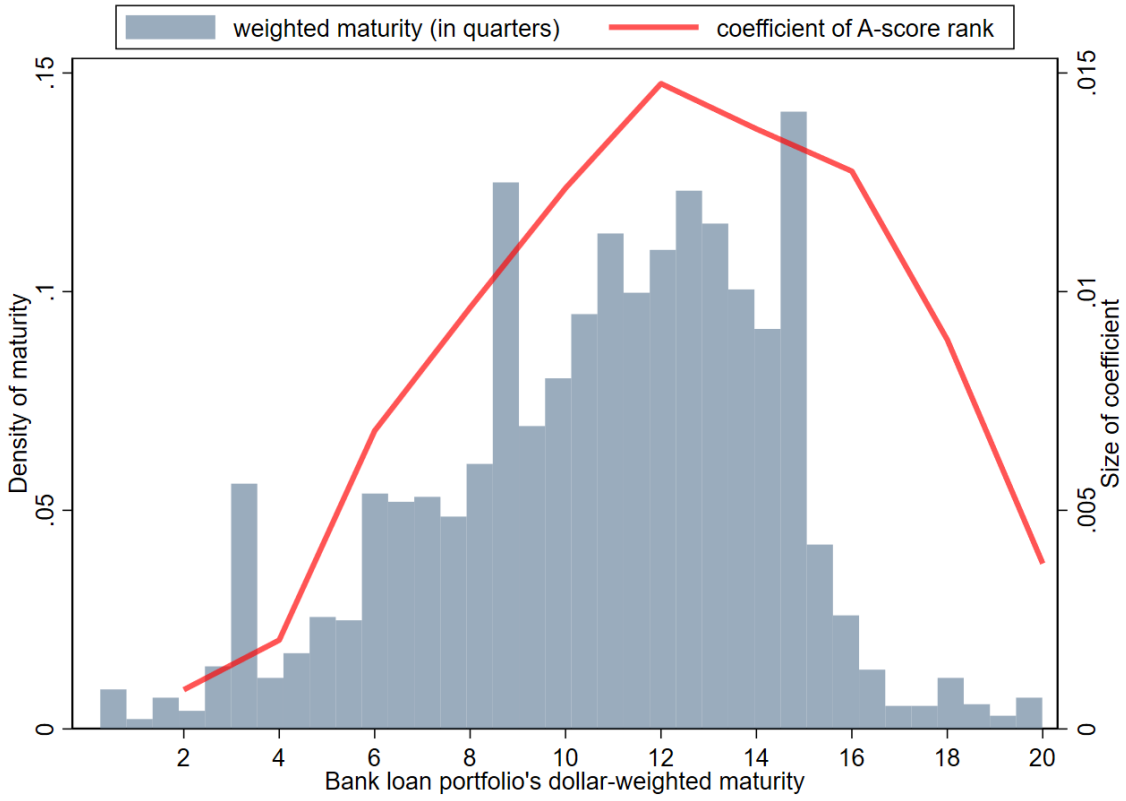


Fig. 4. **Predictive Power of Bank-level A-score by Rank**

Figure 4 plots the coefficient estimates of bank-level A-score by rank in the h-quarter-ahead prediction of future bank loan loss provisions from Table 5 for each prediction horizon h. For brevity, we do not include all coefficient estimates in the figure. Darker lines represent coefficient estimates from higher A-score ranks. We show that the predictive power of bank anomalous lending for loan loss provisions generally come from higher A-score ranks than from lower ranks, suggesting that it is banks with very high A-scores that contribute to the predictive power of A-score.

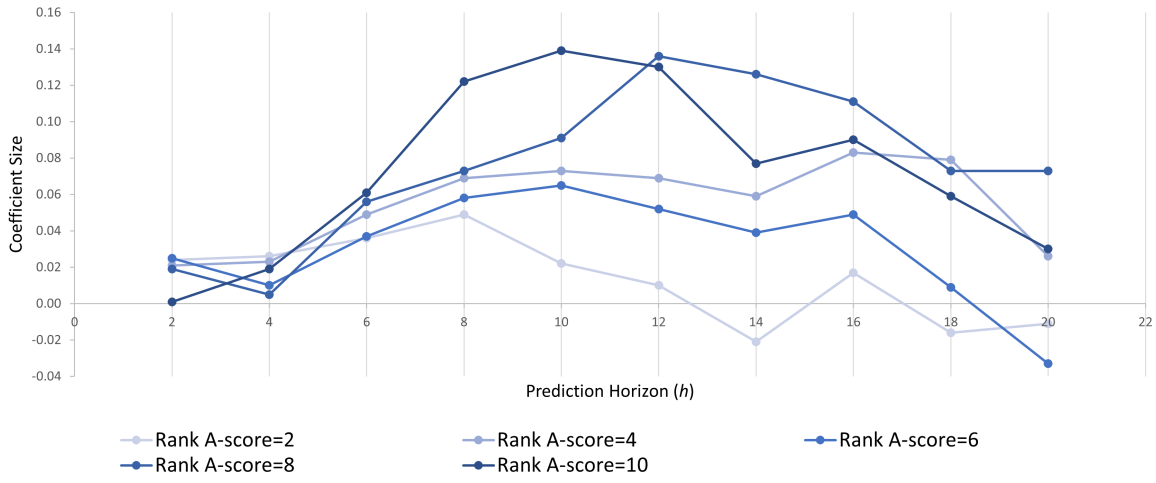


Table 1: **Summary Statistics**

Table 1 presents the summary statistics. The loan sample consists of DealScan syndicated loans originated in the U.S. from 2002 to 2018 with matched lead arranger(s) from FR Y-9C, which is then matched to secondary market trading data based on LIN over the same period. The sample period for the bank-quarter sample is from 1997Q1 to 2020Q1 conditional on the availability of all control variables. Definitions of the variables are provided in Online Appendix OA.1.1. All continuous variables in the cross-sectional analysis are winsorized by year-quarter at the 1st and 99th percentiles.

	N	Mean	Std. Dev.	P10	Median	P90
Loan-level and borrower-level variables						
A-score	34,069	0.442	0.050	0.395	0.427	0.519
Borrower ln(IVOL)	34,069	0.715	0.588	-0.015	0.677	1.483
Borrower default probability	22,868	0.048	0.156	0.000	0.000	0.104
Borrower size	34,069	7.453	2.041	4.848	7.403	10.086
Borrower Tobin's Q	34,069	1.671	0.924	0.990	1.390	2.673
Borrower ROA	34,069	0.005	0.032	-0.016	0.008	0.029
Borrower tangibility	34,069	0.313	0.247	0.039	0.248	0.703
Borrower leverage	34,069	0.345	0.210	0.083	0.330	0.609
Borrower stock return	34,069	1.046	0.244	0.780	1.033	1.307
ln(loan amount)	34,069	5.025	1.660	2.773	5.165	6.975
ln(loan maturity)	34,069	3.648	0.702	2.485	4.025	4.277
All-in-drawn spread	34,069	1.914	1.367	0.375	1.750	3.550
Senior loan	34,069	0.999	0.038	1.000	1.000	1.000
Secured	34,069	0.493	0.500	0.000	0.000	1.000
Performance pricing	34,069	0.443	0.497	0.000	0.000	1.000
Number of lenders	34,069	9.224	9.064	1.000	7.000	20.000
Number of covenants	34,069	1.215	1.350	0.000	1.000	3.000
Lead bank reputation	34,069	0.014	0.031	0.000	0.000	0.049
Lead bank specialization (syndicated loan)	34,069	0.002	0.005	0.000	0.000	0.006
Lead bank specialization (industry)	34,069	0.024	0.054	0.000	0.000	0.063
Has discount	8,642	0.599	0.490	0.000	1.000	1.000
Discount (%)	8,642	4.127	11.102	-1.000	0.387	12.750
Bid-Ask spread (%)	8,642	1.177	1.453	0.312	0.750	2.000
Bank-level variables						
A-score	2,710	0.451	0.050	0.405	0.434	0.551
A-score (loans arranged)	2,710	0.281	0.225	0.000	0.418	0.495
A-score (loans joined)	2,710	0.432	0.102	0.401	0.431	0.551
Loan loss provision	2,710	0.211	0.389	0.002	0.104	0.460
Non-performing loan	2,620	0.790	0.871	0.118	0.504	1.835
Default probability	2,267	0.125	0.165	0.016	0.064	0.286
ln(IVOL)	2,705	0.137	0.499	-0.424	0.050	0.802
Size	2,710	17.277	1.798	15.143	17.084	19.728
Tier 1 capital	2,710	0.100	0.018	0.078	0.100	0.122
Loan size	2,710	0.595	0.169	0.374	0.636	0.763
Deposits	2,710	0.720	0.103	0.585	0.731	0.849
ROA	2,710	0.620	0.839	0.205	0.626	1.271
MTB	2,710	1.797	0.937	0.836	1.611	2.946
Reputation	2,710	0.004	0.013	0.000	0.000	0.004
Specialization in syndicated loan	2,710	0.002	0.004	0.000	0.001	0.005
Specialization in industry	2,710	0.362	0.344	0.051	0.206	1.000
Bank Z-score	2,706	29.929	12.456	17.188	27.285	46.986

Table 2: **Anomalous Loan and Borrower Risk**

Table 2 presents the evidence of loan A-score in predicting future borrower firm risk. Specifically, panel A reports the results where we regress the borrower's idiosyncratic stock return volatility on the quarterly decile rank of loan A-score, borrower characteristics, loan characteristics and lead bank characteristics. Panel B reports the results where the dependent variable is Merton (1974) default probability of the borrower firm. Columns (1) to (4) respectively present results for the dependent variable measured at 0, 4, 8 and 12 quarters after the loan origination. Heteroskedasticity-robust standard errors are clustered at the year-quarter level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Borrower ln(IVOL)

	(1) t	(2) t+4	(3) t+8	(4) t+12
Rank A-score	0.014***	0.010***	0.008***	0.008***
	(5.774)	(4.320)	(3.497)	(3.026)
Borrower size	-0.099***	-0.100***	-0.102***	-0.099***
	(-21.192)	(-19.686)	(-20.305)	(-18.940)
Borrower Tobin's Q	-0.022***	-0.016**	-0.019**	-0.025***
	(-3.453)	(-2.114)	(-2.357)	(-3.071)
Borrower ROA	-1.918***	-1.892***	-1.942***	-1.536***
	(-8.795)	(-8.564)	(-9.436)	(-7.243)
Borrower tangibility	-0.055**	-0.072**	-0.049*	-0.086**
	(-2.042)	(-2.214)	(-1.673)	(-2.452)
Borrower leverage	0.114***	0.104***	0.077***	0.081**
	(4.095)	(4.007)	(2.644)	(2.537)
Borrower stock return	0.032	-0.102***	-0.080***	-0.062**
	(0.789)	(-3.462)	(-2.637)	(-2.576)
ln(loan amount)	0.005	0.001	-0.003	-0.002
	(1.390)	(0.153)	(-0.847)	(-0.387)
ln(loan maturity)	-0.032***	-0.027***	-0.021***	-0.023***
	(-5.123)	(-4.426)	(-3.648)	(-4.194)
All-in-drawn spread	0.094***	0.086***	0.076***	0.062***
	(18.167)	(18.733)	(15.838)	(12.815)
Senior loan	0.196**	0.200**	0.090	0.035
	(2.309)	(2.217)	(1.047)	(0.328)
Secured	0.110***	0.111***	0.107***	0.101***
	(12.453)	(11.941)	(11.242)	(9.290)
Performance pricing	-0.002	0.002	0.004	0.000
	(-0.205)	(0.270)	(0.431)	(0.049)
Number of lenders	0.001	0.001**	0.002***	0.001***
	(0.996)	(2.196)	(3.200)	(2.704)
Number of covenants	-0.002	-0.001	0.002	0.006
	(-0.507)	(-0.213)	(0.374)	(1.146)
Lead bank reputation	-0.928*	0.011	0.442	1.013
	(-1.954)	(0.082)	(1.522)	(1.449)
Lead bank specialization (syndicated loan)	1.839**	0.687	0.456	-1.702
	(2.412)	(0.680)	(0.377)	(-1.129)
Lead bank specialization (industry)	0.058	0.085	0.125	0.156*
	(0.776)	(1.313)	(1.543)	(1.832)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Borrower Industry Fixed Effects	Yes	Yes	Yes	Yes
Loan Purpose Fixed Effects	Yes	Yes	Yes	Yes
Loan Type Fixed Effects	Yes	Yes	Yes	Yes
Observations	34,069	32,805	30,363	27,875
Adjusted R^2	0.573	0.562	0.549	0.526

Table 2: **Continued**

Panel B. Borrower default probability

	(1) t	(2) t+4	(3) t+8	(4) t+12
Rank A-score	0.007***	0.005***	0.003**	0.002*
	(4.948)	(3.809)	(2.455)	(1.827)
Borrower size	-0.001	-0.002	-0.002	-0.002
	(-0.680)	(-1.096)	(-0.826)	(-1.199)
Borrower Tobin's Q	-0.017***	-0.014***	-0.008*	-0.008**
	(-6.551)	(-5.891)	(-1.896)	(-2.314)
Borrower ROA	-0.541***	-0.474***	-0.344***	-0.190***
	(-3.896)	(-4.915)	(-3.969)	(-2.718)
Borrower tangibility	-0.017	-0.039***	-0.009	-0.038**
	(-1.317)	(-2.987)	(-0.622)	(-2.482)
Borrower leverage	0.114***	0.102***	0.078***	0.087***
	(8.849)	(7.214)	(4.824)	(5.062)
Borrower stock return	-0.083***	-0.095***	-0.022*	-0.011
	(-3.406)	(-4.614)	(-1.723)	(-1.137)
ln(loan amount)	0.002	0.003*	-0.002	-0.001
	(1.119)	(1.928)	(-1.612)	(-0.873)
ln(loan maturity)	-0.006**	-0.003	-0.004	-0.007**
	(-2.062)	(-1.172)	(-1.625)	(-2.467)
All-in-drawn spread	0.020***	0.016***	0.012***	0.008***
	(7.013)	(6.977)	(5.355)	(2.871)
Senior loan	-0.059	-0.059	0.009	0.036
	(-0.819)	(-0.875)	(0.153)	(1.124)
Secured	0.011***	0.021***	0.024***	0.021***
	(2.996)	(5.077)	(4.497)	(4.448)
Performance pricing	0.003	0.006	0.000	-0.004
	(0.821)	(1.374)	(0.086)	(-1.095)
Number of lenders	0.000	0.000	0.000	0.000
	(1.524)	(0.327)	(0.170)	(0.727)
Number of covenants	-0.003*	-0.002	-0.004*	0.006**
	(-1.926)	(-1.588)	(-1.968)	(2.245)
Lead bank reputation	0.066	0.029	0.098	0.849*
	(1.037)	(0.400)	(1.160)	(1.693)
Lead bank specialization (syndicated loan)	-0.110	0.618	-0.118	-1.514*
	(-0.424)	(1.515)	(-0.168)	(-1.839)
Lead bank specialization (industry)	-0.004	-0.001	0.037	-0.013
	(-0.164)	(-0.029)	(1.204)	(-0.464)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Borrower Industry Fixed Effects	Yes	Yes	Yes	Yes
Loan Purpose Fixed Effects	Yes	Yes	Yes	Yes
Loan Type Fixed Effects	Yes	Yes	Yes	Yes
Observations	23,518	22,135	20,826	19,273
Adjusted R^2	0.231	0.185	0.164	0.194

Table 3: **Anomalous Loan and Loan Risk**

Table 3 presents the evidence of loan A-score in predicting future secondary loan market trading outcomes. Specifically, columns (1) and (2) investigate whether A-score is related to the likelihood of having a discount on the initial trading day, with and without the depressed loans that are traded at below 90% of face value. Column (3) further studies the size of the discount (in percentage points) and column (4) studies the size of bid-ask spread (in percentage points). We do not include lead arranger fixed effects as many loans have more than one lead arranger, but control for several average lead arranger characteristics. Our results are qualitatively the same if we use the raw A-score instead. Heteroskedasticity-robust standard errors are clustered at the year-quarter level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Has Discount	Has Discount	Discount	Bid-Ask Spread
Rank A-score	0.023***	0.025***	0.029*	0.017***
	(9.874)	(10.001)	(1.846)	(4.080)
ln(loan amount)	-0.035***	-0.017**	-0.134***	-0.094***
	(-3.702)	(-2.311)	(-3.598)	(-9.701)
ln(loan maturity)	-0.096***	-0.093***	0.065	0.054
	(-5.421)	(-5.573)	(0.593)	(1.322)
All-in-drawn spread	0.024***	0.026***	0.047*	0.051***
	(4.167)	(4.016)	(1.782)	(8.591)
Senior loan	0.324***	0.377***	1.993***	0.153**
	(6.535)	(6.938)	(6.649)	(2.216)
Secured	-0.144***	-0.150***	-0.105	-0.079**
	(-8.673)	(-8.762)	(-1.181)	(-2.069)
Performance pricing	0.029	0.031	0.149**	0.076***
	(1.507)	(1.545)	(2.022)	(2.837)
Number of lenders	0.003***	0.003***	0.003	0.001
	(4.153)	(4.000)	(0.874)	(0.675)
Number of covenants	0.005	0.008	-0.005	0.016
	(0.507)	(0.817)	(-0.086)	(1.068)
Lead bank reputation	-0.667	-0.559	1.498	0.802*
	(-1.647)	(-1.169)	(0.690)	(1.898)
Lead bank specialization (syndicated loan)	-4.497	-9.293	-68.729***	-3.785
	(-0.710)	(-1.345)	(-3.202)	(-0.552)
Lead bank specialization (industry)	1.267**	1.386**	3.958	-0.427
	(2.225)	(2.111)	(1.523)	(-0.612)
Excl. Depressed Loans	No	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Borrower Industry Fixed Effects	Yes	Yes	Yes	Yes
Loan Purpose Fixed Effects	Yes	Yes	Yes	Yes
Loan Type Fixed Effects	Yes	Yes	Yes	Yes
Observations	8,642	7,594	7,594	7,594
Adjusted R^2	0.325	0.337	0.242	0.184

Table 4: **Bank Anomalous Lending and Bank Loan Loss Provisions**

Table 4 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions. In all specifications, the bank-level A-score and control variables are measured at time (i.e., end-of-quarter) t, and the dependent variable is the bank's loan loss provisions measured at time $t+h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	0.001	0.002	0.007**	0.010***	0.012***	0.015***	0.014***	0.013**	0.009	0.004
	(0.244)	(0.591)	(2.552)	(3.540)	(3.834)	(3.655)	(2.807)	(2.197)	(1.494)	(0.619)
Size	0.028***	0.025**	0.023***	0.023***	0.015	0.009	0.009	0.010	0.009	0.005
	(2.793)	(2.514)	(2.731)	(2.760)	(1.412)	(0.573)	(0.551)	(0.569)	(0.592)	(0.338)
Tier 1 capital	1.982***	1.327*	0.714	0.111	-0.166	-0.896	-1.118	-1.300	-1.206	-0.978
	(2.756)	(1.894)	(0.907)	(0.142)	(-0.202)	(-1.151)	(-1.327)	(-1.519)	(-1.250)	(-0.993)
Loan size	0.465***	0.526***	0.558***	0.590***	0.602***	0.630***	0.624***	0.627***	0.615***	0.581***
	(4.745)	(5.128)	(5.808)	(6.086)	(6.200)	(6.312)	(6.196)	(6.022)	(5.551)	(4.954)
Deposits	-0.196	-0.303**	-0.338**	-0.343**	-0.504***	-0.578***	-0.579***	-0.511**	-0.488***	-0.474**
	(-1.423)	(-2.035)	(-2.526)	(-2.449)	(-3.027)	(-2.712)	(-2.766)	(-2.468)	(-2.767)	(-2.504)
ROA	-0.053	-0.071**	-0.092***	-0.122***	-0.047***	0.012**	0.020***	0.034***	0.022***	0.021***
	(-1.533)	(-2.203)	(-4.639)	(-4.555)	(-3.001)	(2.286)	(4.740)	(5.353)	(4.468)	(3.756)
MTB	-0.040***	-0.032***	-0.025*	-0.012	-0.024*	-0.031**	-0.022	-0.010	0.011	0.021
	(-4.451)	(-3.127)	(-1.950)	(-0.825)	(-1.756)	(-2.227)	(-1.429)	(-0.635)	(0.671)	(1.176)
Reputation	-0.112	-0.310	-0.683	-0.875*	-1.197**	-1.461**	-1.486**	-1.113*	-0.546	0.154
	(-0.199)	(-0.575)	(-1.414)	(-1.752)	(-2.358)	(-2.600)	(-2.505)	(-1.874)	(-0.941)	(0.274)
Specialization in syndicated loan	0.905	0.791	-0.569	-1.554	-1.557	-3.635	-5.793	-7.814	-9.031	-8.628
	(0.390)	(0.368)	(-0.279)	(-0.548)	(-0.564)	(-1.021)	(-1.407)	(-1.537)	(-1.579)	(-1.269)
Specialization in industry	0.045	0.021	-0.013	-0.024	-0.029	-0.058	-0.065	-0.060	-0.040	-0.022
	(0.978)	(0.483)	(-0.382)	(-0.662)	(-0.707)	(-1.077)	(-1.041)	(-0.922)	(-0.616)	(-0.342)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.573	0.603	0.634	0.583	0.567	0.559	0.561	0.572	0.571

Table 5: **Predictive Power of Anomalous Lending by Rank**

Table 5 presents the h-quarter-ahead prediction results of bank-level A-score by decile rank for future bank loan loss provisions. In all specifications, the bank-level A-score and control variables are measured at time t, and the dependent variable is the bank's loan loss provisions measured at time $t+h$. For brevity, we report only the coefficient estimates of bank A-score ranks but include the same set of controls as in Table 4. Figure 4 provides a visual presentation of this table. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score=2	0.024	0.026	0.036	0.049**	0.022	0.010	-0.021	0.017	-0.016	-0.011
	(0.657)	(0.730)	(1.369)	(2.258)	(0.903)	(0.413)	(-0.947)	(0.890)	(-0.588)	(-0.475)
Rank A-score=3	0.036	0.035	0.076*	0.105***	0.102***	0.089***	0.057**	0.069**	0.053*	0.017
	(0.910)	(0.845)	(1.908)	(3.200)	(2.992)	(3.151)	(2.135)	(2.575)	(1.794)	(0.457)
Rank A-score=4	0.021	0.023	0.049	0.069***	0.073***	0.069**	0.059**	0.083***	0.079**	0.026
	(0.503)	(0.566)	(1.605)	(2.849)	(3.081)	(2.139)	(2.081)	(2.857)	(2.539)	(0.758)
Rank A-score=5	0.007	-0.021	0.052*	0.069**	0.087***	0.083**	0.086**	0.087**	0.066*	0.014
	(0.173)	(-0.503)	(1.847)	(2.350)	(2.736)	(2.565)	(2.523)	(2.598)	(1.906)	(0.393)
Rank A-score=6	0.025	0.010	0.037	0.058**	0.065***	0.052*	0.039	0.049	0.009	-0.033
	(0.614)	(0.241)	(1.347)	(2.292)	(2.685)	(1.882)	(1.486)	(1.517)	(0.314)	(-0.949)
Rank A-score=7	0.054	0.050	0.033	0.088***	0.124***	0.127**	0.080**	0.096**	0.025	-0.007
	(1.559)	(1.169)	(1.249)	(3.631)	(3.664)	(2.643)	(2.345)	(2.516)	(0.920)	(-0.193)
Rank A-score=8	0.019	0.005	0.056*	0.073***	0.091***	0.136***	0.126**	0.111**	0.073*	0.073
	(0.379)	(0.118)	(1.802)	(2.850)	(4.368)	(3.847)	(2.292)	(2.432)	(1.804)	(1.024)
Rank A-score=9	0.015	0.063	0.150***	0.149***	0.115***	0.111***	0.111**	0.143**	0.136	0.039
	(0.314)	(1.391)	(3.460)	(3.969)	(3.856)	(3.430)	(2.526)	(2.113)	(1.551)	(0.768)
Rank A-score=10	0.001	0.019	0.061**	0.122***	0.139**	0.130***	0.077*	0.090*	0.059	0.030
	(0.020)	(0.501)	(2.393)	(3.155)	(2.244)	(2.737)	(1.788)	(1.980)	(1.461)	(0.708)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.575	0.608	0.637	0.586	0.569	0.561	0.563	0.577	0.573

Table 6: **Incremental R^2 and Prediction Exercise**

Table 6 presents the 3-year-ahead prediction results of bank-level A-score for future bank loan loss provisions. In Panel A, we report the in-sample predictions and R^2. Panel B reports the RMSEs of out-of-sample predictions. Specifically, we use a rolling 3-year window to estimate the model specified by Equation 2, with which we then produce four out-of-sample 3-year ahead predictions of bank loan loss provisions using the realized values of the independent variables (in Equation 3) in each of the next four quarters as inputs, respectively. We then pool all prediction errors together to compute the root-mean-squared-error (RMSE). Heteroskedasticity-robust standard errors are clustered at the bank level. We include the same set of control variables and fixed effects as in Table 4. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Controls only	(2) Z-score	(3) A-score	(4) Both
Panel A: In-sample predictions				
Rank A-score			0.015***	0.015***
			(3.645)	(3.515)
Bank Z-score		0.000		0.001
		(0.294)		(0.407)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,140	2,140	2,140	2,140
R^2	0.579	0.579	0.587	0.587
Panel B: Out-of-sample predictions				
RMSE	0.424	0.417	0.411	0.404

Table 7: **Predicting Bank Loan Loss Provisions: A Decomposition of A-score**

Table 7 presents the baseline h-quarter-ahead prediction results of bank-level quarterly A-scores estimated using the four categories of DealScan loan attributes. In all specifications, the bank-level A-scores and control variables are measured at time t. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score (loan covenant)	0.010***	0.008**	0.010***	0.012***	0.012***	0.012***	0.011***	0.011***	0.010**	0.007
	(2.715)	(2.400)	(3.580)	(4.151)	(3.456)	(3.440)	(2.823)	(2.802)	(2.410)	(1.502)
Rank A-score (loan pricing)	-0.001	-0.001	0.000	0.003	0.008*	0.009*	0.008	0.006	0.003	0.001
	(-0.386)	(-0.463)	(0.043)	(0.831)	(1.919)	(1.887)	(1.645)	(1.127)	(0.611)	(0.156)
Rank A-score (loan nonpricing)	-0.005	-0.002	-0.002	-0.002	-0.002	0.001	0.001	0.002	-0.001	-0.003
	(-1.410)	(-0.413)	(-0.625)	(-0.475)	(-0.292)	(0.208)	(0.184)	(0.233)	(-0.158)	(-0.557)
Rank A-score (borrower)	-0.007*	-0.006*	-0.001	-0.002	-0.005	-0.006	-0.004	-0.004	-0.003	-0.003
	(-2.007)	(-1.818)	(-0.283)	(-0.507)	(-0.976)	(-1.521)	(-1.220)	(-0.942)	(-0.916)	(-0.900)
Size	0.029***	0.024**	0.024***	0.024***	0.014	0.006	0.007	0.009	0.010	0.006
	(2.809)	(2.463)	(2.812)	(2.788)	(1.189)	(0.385)	(0.419)	(0.493)	(0.586)	(0.386)
Tier 1 capital	1.841**	1.168*	0.586	-0.037	-0.292	-0.973	-1.190	-1.376	-1.303	-1.084
	(2.628)	(1.725)	(0.751)	(-0.048)	(-0.361)	(-1.234)	(-1.411)	(-1.633)	(-1.395)	(-1.148)
Loan size	0.481***	0.540***	0.580***	0.627***	0.653***	0.677***	0.668***	0.664***	0.640***	0.596***
	(4.971)	(5.313)	(6.042)	(6.237)	(6.316)	(6.354)	(6.280)	(6.221)	(5.892)	(5.073)
Deposits	-0.176	-0.295*	-0.322**	-0.327**	-0.504***	-0.596**	-0.590**	-0.507**	-0.469***	-0.451**
	(-1.242)	(-1.896)	(-2.383)	(-2.300)	(-2.741)	(-2.534)	(-2.664)	(-2.446)	(-2.745)	(-2.529)
ROA	-0.052	-0.070**	-0.091***	-0.122***	-0.046***	0.012**	0.020***	0.035***	0.022***	0.020***
	(-1.536)	(-2.210)	(-4.625)	(-4.568)	(-3.166)	(2.281)	(4.623)	(5.277)	(4.334)	(3.583)
MTB	-0.037***	-0.031***	-0.024*	-0.011	-0.024*	-0.031**	-0.022	-0.010	0.012	0.023
	(-4.231)	(-2.894)	(-1.874)	(-0.791)	(-1.739)	(-2.197)	(-1.416)	(-0.661)	(0.674)	(1.247)
Reputation	0.385	0.069	-0.361	-0.502	-0.692	-0.942	-1.083*	-0.769	-0.167	0.518
	(0.665)	(0.121)	(-0.694)	(-0.940)	(-1.324)	(-1.650)	(-1.789)	(-1.211)	(-0.295)	(0.961)
Specialization in syndicated loan	1.106	0.851	-0.573	-1.679	-1.658	-3.700	-5.936	-8.003	-8.980	-8.481
	(0.460)	(0.393)	(-0.292)	(-0.610)	(-0.634)	(-1.107)	(-1.477)	(-1.593)	(-1.622)	(-1.284)
Specialization in industry	0.027	0.008	-0.027	-0.042	-0.041	-0.061	-0.070	-0.066	-0.048	-0.032
	(0.665)	(0.216)	(-0.867)	(-1.239)	(-1.120)	(-1.276)	(-1.211)	(-1.043)	(-0.791)	(-0.534)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.561	0.576	0.605	0.637	0.588	0.570	0.561	0.563	0.573	0.571

Table 8: **Loans Joined vs Loan Lead-arranged**

Table 8 re-estimates the baseline h-quarter ahead prediction model (Table 4) by replacing the bank-level A-score rank with 1) A-score rank based on loans joined by the bank (excluding loans arranged) and 2) A-score rank based on loans lead-arranged by the bank. In all specifications, bank-level A-score and control variables are measured at time t, and the dependent variable is the bank's loan loss provisions measured at time $t+h$. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score loans joined	0.005	0.005	0.009***	0.010***	0.011***	0.013***	0.012***	0.012**	0.005	-0.001
	(1.498)	(1.581)	(3.306)	(3.588)	(3.686)	(3.782)	(3.074)	(2.169)	(0.972)	(-0.175)
Rank A-score loans arranged	-0.001	-0.000	-0.000	-0.002	-0.001	-0.001	0.000	0.000	0.001	0.001
	(-0.180)	(-0.052)	(-0.127)	(-0.464)	(-0.396)	(-0.233)	(0.015)	(0.004)	(0.169)	(0.196)
Size	0.029***	0.026**	0.024***	0.025***	0.017	0.011	0.011	0.010	0.009	0.005
	(2.768)	(2.512)	(2.924)	(3.065)	(1.474)	(0.654)	(0.599)	(0.566)	(0.521)	(0.281)
Tier 1 capital	1.967***	1.297*	0.662	0.062	-0.217	-0.987	-1.256	-1.442	-1.254	-0.966
	(2.741)	(1.872)	(0.851)	(0.081)	(-0.270)	(-1.281)	(-1.510)	(-1.665)	(-1.277)	(-0.972)
Loan size	0.465***	0.527***	0.562***	0.591***	0.603***	0.635***	0.634***	0.636***	0.618***	0.580***
	(4.722)	(5.120)	(5.867)	(6.145)	(6.231)	(6.343)	(6.239)	(5.964)	(5.534)	(4.941)
Deposits	-0.184	-0.291*	-0.317**	-0.320**	-0.485***	-0.551**	-0.560**	-0.502**	-0.495***	-0.488**
	(-1.292)	(-1.903)	(-2.374)	(-2.318)	(-2.889)	(-2.542)	(-2.624)	(-2.405)	(-2.772)	(-2.603)
ROA	-0.052	-0.070**	-0.091***	-0.122***	-0.047***	0.012**	0.020***	0.034***	0.022***	0.020***
	(-1.519)	(-2.191)	(-4.563)	(-4.465)	(-2.924)	(2.279)	(4.674)	(5.289)	(4.114)	(3.615)
MTB	-0.041***	-0.033***	-0.026*	-0.012	-0.024*	-0.031**	-0.022	-0.011	0.012	0.023
	(-4.487)	(-3.152)	(-1.906)	(-0.862)	(-1.772)	(-2.261)	(-1.468)	(-0.703)	(0.682)	(1.247)
Reputation	-0.257	-0.402	-0.641	-0.697	-0.910*	-1.119**	-1.206**	-0.887	-0.266	0.381
	(-0.490)	(-0.770)	(-1.302)	(-1.356)	(-1.750)	(-2.015)	(-2.093)	(-1.547)	(-0.474)	(0.687)
Specialization in syndicated loan	0.455	0.440	-0.838	-1.661	-1.522	-3.538	-5.700	-7.743	-8.702	-8.209
	(0.209)	(0.212)	(-0.404)	(-0.585)	(-0.569)	(-1.023)	(-1.390)	(-1.497)	(-1.528)	(-1.219)
Specialization in industry	0.041	0.019	-0.009	-0.021	-0.021	-0.045	-0.051	-0.051	-0.024	-0.007
	(0.953)	(0.458)	(-0.284)	(-0.605)	(-0.603)	(-0.983)	(-1.011)	(-0.945)	(-0.436)	(-0.124)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.558	0.574	0.604	0.634	0.582	0.565	0.557	0.561	0.570	0.570

Table 9: **Bank Performance, Anomalous Lending and Loan Loss Provisions**

Table 9 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, as moderated by the bank's recent stock market or accounting performance. In all specifications, we include bank-level A-score rank, bank performance measure and their interaction, as well as the same set of control variables as in Table 4 measured at time t. The dependent variable is the bank's loan loss provisions measured at time $t + h$. For simplicity, we do not report the coefficient estimates of the control variables. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Market performance: buy-and-hold abnormal return										
Rank A-score	0.001	0.002	0.008***	0.011***	0.014***	0.016***	0.015***	0.013**	0.008	0.004
	(0.160)	(0.633)	(2.814)	(3.950)	(4.163)	(3.531)	(2.884)	(2.175)	(1.468)	(0.608)
Rank A-score × BHAR	0.003	-0.003	-0.016	-0.027***	-0.021**	-0.024*	-0.019**	0.005	0.012	0.005
	(0.231)	(-0.250)	(-1.485)	(-3.523)	(-2.279)	(-1.807)	(-2.509)	(0.764)	(1.381)	(0.646)
BHAR	-0.155**	-0.160**	-0.074	0.025	0.006	0.030	0.026	-0.078	-0.069	-0.043
	(-2.044)	(-2.165)	(-1.126)	(0.496)	(0.106)	(0.441)	(0.538)	(-1.428)	(-1.098)	(-0.687)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.562	0.583	0.612	0.640	0.588	0.571	0.561	0.562	0.572	0.571
Accounting performance: earnings surprise										
Rank A-score	0.002	0.002	0.008***	0.011***	0.014***	0.017***	0.016***	0.016**	0.012*	0.007
	(0.394)	(0.642)	(2.715)	(3.905)	(4.294)	(4.075)	(3.041)	(2.473)	(1.852)	(1.101)
Rank A-score × Earnings surprise	0.019	0.003	0.008	0.059	-0.057**	-0.024	-0.035***	-0.023***	0.000	0.014
	(0.463)	(0.041)	(0.422)	(0.848)	(-2.337)	(-0.445)	(-3.835)	(-3.357)	(0.057)	(1.588)
Earnings surprise	-0.340	-0.116	0.012	-0.180	0.353**	0.061	0.185***	0.105**	-0.008	-0.092*
	(-1.179)	(-0.216)	(0.092)	(-0.434)	(2.315)	(0.180)	(2.972)	(2.348)	(-0.167)	(-1.698)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,482	2,374	2,307	2,231	2,158	2,089	2,013	1,942	1,882	1,817
Adjusted R^2	0.563	0.576	0.604	0.637	0.584	0.570	0.564	0.569	0.582	0.583

Table 10: **Bank Complexity, Anomalous Lending and Loan Loss Provisions**

Table 10 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, as moderated by the bank's complexity, measured as the ratio of the number of nonmissing items to the number of total items in FR Y-9C orthogonalized to bank size. In all specifications, we include bank-level A-score rank, bank complexity measure and their interaction, as well as the same set of control variables as in Table 4 measured at time t. The dependent variable is the bank's loan loss provisions measured at time $t+h$. For simplicity, we do not report the coefficient estimates of the control variables. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h=2	h=4	h=6	h=8	h=10	h=12	h=14	h=16	h=18	h=20
Rank A-score	0.000	0.001	0.006**	0.009***	0.011***	0.014***	0.013***	0.012**	0.008	0.003
	(0.127)	(0.438)	(2.483)	(3.283)	(3.332)	(3.296)	(2.697)	(2.127)	(1.474)	(0.581)
Rank A-score × Complexity	0.017	0.073	0.136*	0.185**	0.272***	0.285***	0.297***	0.270**	0.262**	0.168*
	(0.273)	(0.968)	(1.918)	(2.408)	(3.918)	(3.525)	(3.311)	(2.437)	(2.385)	(1.729)
Complexity	-0.748	-1.043**	-1.068**	-1.110**	-1.542***	-1.567***	-1.701***	-1.333**	-1.054	-0.514
	(-1.403)	(-2.326)	(-2.531)	(-2.651)	(-3.484)	(-3.825)	(-3.468)	(-2.255)	(-1.563)	(-0.837)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.558	0.575	0.605	0.636	0.589	0.573	0.565	0.567	0.578	0.573

Table 11: **Bank Anomalous Lending and Alternative Bank Risk Measures**

Table 11 presents the baseline h-quarter-ahead prediction results of bank-level A-score for alternative bank risk measures. In all specifications, bank-level A-score and control variables are measured at time t and the dependent variable is measured at time $t+h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Dependent: nonperforming loan										
Rank A-score	0.013	0.014	0.020*	0.023**	0.030***	0.033***	0.028***	0.026**	0.025**	0.019
	(1.602)	(1.583)	(1.941)	(2.120)	(2.733)	(3.172)	(2.754)	(2.514)	(2.314)	(1.450)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,477	2,380	2,333	2,270	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.620	0.637	0.631	0.628	0.627	0.629	0.623	0.621	0.611	0.596
Dependent: Merton default probability										
Rank A-score	0.003*	0.004**	0.006***	0.006***	0.006***	0.007***	0.008***	0.008***	0.006***	0.005**
	(1.690)	(2.658)	(2.851)	(2.782)	(3.343)	(3.674)	(4.012)	(4.229)	(3.157)	(2.464)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,092	1,954	1,882	1,803	1,743	1,689	1,616	1,546	1,481	1,411
Adjusted R^2	0.733	0.742	0.745	0.741	0.736	0.726	0.726	0.748	0.738	0.723
Dependent: modified default probability										
Rank A-score	0.002	0.002	0.002	0.003	0.003	0.004*	0.004**	0.004**	0.003	0.001
	(1.032)	(1.072)	(1.122)	(1.221)	(1.485)	(1.815)	(2.152)	(2.233)	(1.620)	(0.590)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,092	1,954	1,882	1,803	1,743	1,689	1,616	1,546	1,481	1,411
Adjusted R^2	0.702	0.709	0.693	0.677	0.669	0.652	0.646	0.660	0.653	0.629
Dependent: ln(IVOL)										
Rank A-score	0.012**	0.013***	0.014***	0.014***	0.016***	0.021***	0.017***	0.017***	0.017***	0.012***
	(2.663)	(3.125)	(3.168)	(2.941)	(3.649)	(5.530)	(4.515)	(4.217)	(4.010)	(2.889)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,440	2,366	2,281	2,200	2,125	2,045	1,972	1,902	1,827
Adjusted R^2	0.693	0.705	0.703	0.697	0.695	0.702	0.692	0.695	0.691	0.684

Table 12: **Impact of Stress Test Failure on Bank Anomalous Lending**

Table 12 examines the impact of bank stress test failures on bank A-score as measured at each year's end. The stacked-cohort difference-in-differences (DiD) estimations are conducted using the CCAR test results from 2012 to 2019, where the sample consists of all banks that have ever participated in any of the CCAR tests. The treatment events are the banks' first CCAR test failures. In each cohort, Treat equals 1 for the bank that failed the CCAR test and 0 for banks that have never or not yet failed any CCAR test by that year, with a five-year event window from two years before the failure of the treated bank til two years afterwards. Further, Post equals 1 (0) for all years after (before) the treatment year in each cohort, and time dummies d_j equals 1 for the year that is j year(s) relative to the focal CCAR-test year. Definitions of the variables are provided in Online Appendix OA.1.1. An overview of stress test outcomes is provided in Online Appendix OA.1.12. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Standard DiD

	(1)	(2)	(3)	(4)
Treat × Post	-0.004**	-0.004**	-0.005**	-0.005**
	(-2.525)	(-2.412)	(-2.669)	(-2.644)
Treat	0.001	0.001		
	(0.939)	(0.939)		
Post	-0.000	-0.000*		
	(-0.919)	(-2.009)		
Bank Fixed Effects	Yes	Yes	No	No
Year Fixed Effects	Yes	Yes	No	No
Cohort-Bank Fixed Effects	No	No	Yes	Yes
Cohort-Year Fixed Effects	No	No	Yes	Yes
Standard Error Clustering	Bank	Bank & Cohort	Bank	Bank & Cohort
Observations	1,300	1,300	1,280	1,280
Adjusted R^2	0.988	0.988	0.987	0.987

Panel B. Dynamic DiD

	(1)	(2)	(3)	(4)
Treat × d_{-1}	-0.003	-0.003	-0.004	-0.004
	(-0.884)	(-0.832)	(-1.100)	(-1.067)
Treat × d_0	-0.006**	-0.006**	-0.007***	-0.007***
	(-2.726)	(-2.584)	(-3.060)	(-3.027)
Treat × d_1	-0.006	-0.006	-0.007**	-0.007*
	(-1.628)	(-1.531)	(-2.136)	(-2.095)
Treat × d_2	-0.004	-0.004	-0.006**	-0.006*
	(-1.427)	(-1.344)	(-2.103)	(-2.070)
Treat and Time Dummies	Yes	Yes	No	No
Bank Fixed Effects	Yes	Yes	No	No
Year Fixed Effects	Yes	Yes	No	No
Cohort-Bank Fixed Effects	No	No	Yes	Yes
Cohort-Year Fixed Effects	No	No	Yes	Yes
Standard Error Clustering	Bank	Bank & Cohort	Bank	Bank & Cohort
Observations	1,300	1,300	1,280	1,280
Adjusted R^2	0.988	0.988	0.987	0.987

Online Appendix

OA.1. Variation definition and additional results

Table OA.1.1: **Variable Definition**

Variable	Definition	Source
Loan-level variables		
A-score	Loan-level anomaly score computed by applying the Isolation Forest algorithm on DealScan. The detailed estimation process is described in Section 4 and Online Appendix OA.2.	Our computation
ln(loan amount)	The natural logarithm of the loan dollar amount.	DealScan
ln(loan maturity)	The natural logarithm of the loan maturity in months.	DealScan
All-in-drawn spread	The all-in-drawn spread (percentage points) of the loans.	DealScan
Senior loan	An indicator variable that takes the value of one, and zero otherwise, if the loan is a senior loan.	DealScan
Secured	An indicator variable that takes the value of one, and zero otherwise, if the loan is secured.	DealScan
Performance pricing	An indicator variable that takes the value of one, and zero otherwise, if the loan has performance pricing.	DealScan
Number of lenders	The number of lenders in the loan syndicate.	DealScan
Number of covenants	The number of financial covenants in the loan contract.	DealScan
Lead bank reputation	The average of all lead arrangers' reputation, defined as the total amount of syndicated loans lead-arranged by a bank divided by the total amount of all syndicated loans in the past 12 months.	DealScan
Lead bank specialization (syndicated loan)	The average of all lead arrangers' specialization in syndicated loan, defined as the total dollar amount of the syndicated loans the bank participates over the past 12 months normalized by the total loans (BHCK2122).	DealScan, FR Y-9C
Lead bank specialization (industry)	The average of all lead arrangers' specialization in industry, defined as the borrower's industry concentration of the bank's syndicated loan portfolio over the past 12 months, measured by the HHI index based on the borrowers' 2-digit SIC codes and the loan amounts.	DealScan
Borrower-level variables		
Borrower ln(IVOL)	Idiosyncratic volatility of borrower stock returns, measured as the natural logarithm of the standard deviation of the residuals from the Fama-French three-factor model estimated for each year-quarter.	CRSP, Kenneth R. French Data Library
Borrower default probability	The probability of default estimated using the Merton (1974) model via a KMV iterative approach.	Our computation
Borrower size	The natural logarithm of total assets (ATQ).	Compustat
Borrower Tobin's Q	The firm's Tobin's Q, computed as (ATQ+(CSHOQ*PRCCQ)-CEQQ)/ATQ.	Compustat
Borrower ROA	Total income (IBQ) normalized by total assets (ATQ).	Compustat
Borrower tangibility	Property, Plant and Equipment (PPENTQ) normalized by total assets (ATQ).	Compustat
Borrower leverage	Total debt (DLTTQ + DLCQ) normalized by total assets (ATQ).	Compustat
Borrower stock return	The quarterly buy-and-hold stock return.	CRSP

Table OA.1.1: **Continued**

Variable	Definition	Source
Bank-level variables		
Loan loss provisions	The ratio of loan loss provisions (BHCK4230) to total assets (BHCK2170) in percentage points.	FR Y-9C
Nonperforming loans	The ratio of nonperforming loans (sum of BHCK5525 and BHCK5526 from 1990Q3 to 2017Q4 and sum of BHCK1403 and BHCK1407 after 2017) to total assets (BHCK2170) in percentage points.	FR Y-9C
Merton default probability	The probability of default estimated using the Merton (1974) model via a KMV iterative approach.	Our computation
Modified default probability	The modified probability of default using the Nagel and Purnanandam (2020) model.	Nagel and Purnanandam (2020)
ln(IVOL)	Idiosyncratic volatility of bank stock returns, measured as the natural logarithm of the standard deviation of the residuals from the Fama-French three-factor model estimated for each year-quarter.	CRSP, Kenneth R. French Data Library
Size	The natural logarithm of total assets (BHCK2170).	FR Y-9C
Tier 1 capital	The ratio of a bank's Tier 1 capital (BHCK3792 from 1997 to 2014 and BHCA3792 since 2014) to total assets (BHCK2170).	FR Y-9C
Loan size	Total loans (BHCK2122) normalized by total assets (BHCK2170).	FR Y-9C
Deposits	Total deposits (sum of BHDM6631, BHDM6636, BHFN6631 and BHFN6636) normalized by total assets (BHCK2170).	FR Y-9C
ROA	Net income (BHCK4340) normalized by total assets (BHCK2170) times 100.	FR Y-9C
MTB	The ratio of a bank's market capitalization to its book value of total equity.	FR Y-9C, CRSP
Reputation	The total amount of syndicated loans lead-arranged by a bank divided by the total amount of all syndicated loans in the past 12 months.	DealScan
Specialization in syndicated loan	The total dollar amount of the syndicated loans a bank participates over the past 12 months normalized by the total loans (BHCK2122).	DealScan, FR Y-9C
Specialization in industry	Borrower's industry concentration of a bank's syndicated loan portfolio over the past 12 months, measured by the HHI index based on the borrowers' 2-digit SIC codes and the loan amounts.	DealScan
Bank Z-score	The bank's Z-score, defined as the sum of ROA and equity ratio divided by the standard deviation of bank's ROA in the past five years.	FR Y-9C
BHAR	The buy-and-hold abnormal return in the past year computed using the market model.	CRSP
Earnings surprise	The actual earnings minus the average analysts' forecast, normalized by the stock price.	I/B/E/S, CRSP
Complexity	The ratio of the number of non-missing items to the total number of items in FR Y-9C in each year-quarter, orthogonalized to bank size.	FR Y-9C
Nonperforming loans to individuals	The ratio of nonperforming loans to individuals (sum of BHCK5384, BHCK5385, BHCK5387 and BHCK5388 before 2001, sum of BHCKb576, BHCKb577, BHCKb579 and BHCKb580 from 2001 to 2010, and sum of BHCKb576, BHCKb577, BHCKk214, BHCKk215, BHCKk217 and BHCKk218 after 2010) to total assets (BHCK2170) in percentage points.	FR Y-9C
Nonperforming loans to firms	The ratio of nonperforming loans commercial and industrial (sum of BHCK1607 and BHCK1608) and nonperforming loans secured by real estate (sum of BHCK1422 and BHCK1423 before 2001, sum of BHCK2769, BHCK3492, BHCK3494, BHCK3495, BHCK5399, BHCK5400, BHCK5402, BHCK5403, BHCK3500, BHCK3501, BHCK3503, BHCK3504 from 2001 to 2007, and sum of BHCKC229, BHCKC237, BHCKF174, BHCKF175, BHCKF176, BHCKF177, BHCKF180, BHCKF181, BHCKF182, BHCKF183 after 2007) to total assets (BHCK2170) in percentage points.	FR Y-9C

Table OA.1.2: **Banks with High and Low Anomalous Lending**

Table OA.1.2 lists the top and bottom five banks in the sample with the highest and lowest anomalous lending activities measured by bank-level A-score. Specifically, we require a bank to have at least 10 years or 40 quarters of data to be included. Top (%) (Bottom (%)) refers to the proportion of the bank's quarterly A-score ranks that are above (below) median, relative to the total number of its A-score estimates.

Bank	Average A-score Rank	Top (%)	Bottom (%)
Top five			
CIT Group Inc	9.15	100.00	0.00
Ocean Bank	8.53	97.67	2.33
Goldman Sachs & Co	8.24	100.00	0.00
Morgan Stanley	8.16	100.00	0.00
Bank of New York Mellon	7.20	92.16	7.84
Bottom five			
KeyBank	3.89	11.11	88.89
Regions Bank	3.73	6.35	93.65
Banco Bilbao Vizcaya Argentaria SA [BBVA]	3.59	9.91	90.09
FirstMerit Bank NA	2.96	10.98	89.02
Webster Bank NA	2.95	15.63	84.38

Table OA.1.3: **Anomalous Loan and Borrower Risk: Term Loans**

Table OA.1.3 presents the evidence of loan A-score in predicting future borrower firm risk as moderated by loan type. Specifically, panel A reports the results where we regress the borrower's idiosyncratic stock return volatility on the quarterly decile rank of loan A-score, borrower characteristics, loan characteristics and lead bank characteristics. Panel B reports the results where the dependent variable is Merton (1974) default probability of the borrower firm. Columns (1) to (4) respectively present results for the dependent variable measured at 0, 4, 8 and 12 quarters after the loan origination. Heteroskedasticity-robust standard errors are clustered at the year-quarter level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Borrower's ln(IVOL)

	(1)	(2)	(3)	(4)
	t	t+4	t+8	t+12
Rank A-score	0.014***	0.010***	0.010***	0.009***
	(5.670)	(4.019)	(4.180)	(3.190)
Rank A-score × Term loan	0.001	-0.001	-0.005	-0.002
	(0.532)	(-0.219)	(-1.540)	(-0.802)
Control Variables	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Borrower Industry Fixed Effects	Yes	Yes	Yes	Yes
Loan Purpose Fixed Effects	Yes	Yes	Yes	Yes
Loan Type Fixed Effects	Yes	Yes	Yes	Yes
Observations	34,069	32,805	30,363	27,875
Adjusted R^2	0.573	0.562	0.549	0.526

Panel B. Borrower's default probability

	(1)	(2)	(3)	(4)
	t	t+4	t+8	t+12
Rank A-score	0.640***	0.424***	0.248**	0.131
	(4.438)	(3.218)	(2.075)	(1.183)
Rank A-score × Term loan	0.240*	0.222	0.061	0.258
	(1.833)	(1.456)	(0.441)	(1.579)
Control Variables	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes
Borrower Industry Fixed Effects	Yes	Yes	Yes	Yes
Loan Purpose Fixed Effects	Yes	Yes	Yes	Yes
Loan Type Fixed Effects	Yes	Yes	Yes	Yes
Observations	23,518	22,135	20,826	19,273
Adjusted R^2	0.231	0.185	0.164	0.195

Table OA.1.4: **Bank Anomalous Lending and Bank Loan Loss Provisions with Bank Fixed Effect**

Table OA.1.4 presents the baseline h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, same as in Table 4, additionally controlling for bank fixed effects. In all specifications, bank-level A-score and control variables are measured at time t and the dependent variable is the bank's loan loss provisions measured at time $t + h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	-0.003	-0.001	0.005	0.009***	0.010***	0.013***	0.012***	0.011**	0.007	-0.001
	(-0.498)	(-0.157)	(1.422)	(2.705)	(2.729)	(3.810)	(3.060)	(2.547)	(1.302)	(-0.105)
Size	0.026	0.025	0.038	0.059*	0.057	0.063	0.061	0.053	0.045	0.039
	(0.849)	(0.759)	(1.160)	(1.816)	(1.595)	(1.539)	(1.545)	(1.368)	(1.086)	(0.926)
Tier 1 capital	0.449	-0.394	-1.556	-2.797***	-2.791**	-3.586***	-3.473***	-3.608***	-3.092**	-2.239*
	(0.487)	(-0.413)	(-1.466)	(-2.775)	(-2.281)	(-3.124)	(-2.765)	(-3.180)	(-2.406)	(-1.885)
Loan size	0.677**	0.867***	0.819***	0.714***	0.658***	0.594***	0.467**	0.416*	0.242	0.069
	(2.352)	(2.702)	(3.709)	(4.373)	(3.911)	(3.610)	(2.428)	(2.006)	(1.103)	(0.283)
Deposits	-0.385	-0.558**	-0.631**	-0.596**	-0.740**	-0.730**	-0.740**	-0.510	-0.436*	-0.262
	(-1.599)	(-2.129)	(-2.436)	(-2.149)	(-2.452)	(-2.084)	(-2.219)	(-1.630)	(-1.895)	(-1.168)
ROA	-0.043	-0.063*	-0.092***	-0.127***	-0.046***	0.017***	0.026***	0.042***	0.031***	0.031***
	(-1.297)	(-2.006)	(-4.964)	(-4.886)	(-2.922)	(3.512)	(6.539)	(7.706)	(9.109)	(5.367)
MTB	-0.077***	-0.069***	-0.058***	-0.039***	-0.042***	-0.042***	-0.026	-0.008	0.021	0.035**
	(-5.033)	(-5.178)	(-4.893)	(-4.162)	(-3.600)	(-3.036)	(-1.622)	(-0.567)	(1.457)	(2.645)
Reputation	-2.110	-1.788	-2.174*	-1.947**	-1.824*	-1.952	-2.827	-2.484	-2.249	-2.147
	(-1.173)	(-1.122)	(-1.998)	(-2.189)	(-1.805)	(-1.468)	(-1.628)	(-1.450)	(-1.305)	(-1.362)
Specialization in syndicated loan	0.826	0.997	0.142	-0.122	0.300	-1.137	-2.990	-4.315	-5.150	-4.409
	(0.350)	(0.502)	(0.097)	(-0.073)	(0.195)	(-0.577)	(-1.282)	(-1.439)	(-1.556)	(-1.065)
Specialization in industry	0.064	0.059	0.035	0.033	0.041	-0.005	-0.013	-0.012	-0.008	-0.011
	(1.142)	(0.975)	(0.773)	(0.685)	(0.709)	(-0.072)	(-0.176)	(-0.191)	(-0.141)	(-0.230)
Bank and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,442	2,373	2,291	2,212	2,143	2,067	1,998	1,931	1,860
Adjusted R^2	0.602	0.616	0.641	0.670	0.618	0.606	0.598	0.603	0.615	0.619

Table OA.1.5: **Bank Anomalous Lending and Bank Loan Loss Provisions with Double Standard Errors Clustering**

Table OA.1.5 presents the baseline h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, same as in Table 4. In all specifications, bank-level A-score and control variables are measured at time t and the dependent variable is the bank's loan loss provisions measured at time $t + h$. Heteroskedasticity-robust standard errors are clustered at the year-quarter and bank levels. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	0.001	0.002	0.007**	0.010***	0.012***	0.015***	0.014**	0.013**	0.009	0.004
	(0.259)	(0.609)	(2.387)	(2.955)	(2.859)	(2.908)	(2.425)	(2.012)	(1.372)	(0.596)
Size	0.028***	0.025**	0.023***	0.023***	0.015	0.009	0.009	0.010	0.009	0.005
	(2.877)	(2.673)	(2.973)	(2.998)	(1.500)	(0.613)	(0.588)	(0.607)	(0.643)	(0.370)
Tier 1 capital	1.982**	1.327*	0.714	0.111	-0.166	-0.896	-1.118	-1.300	-1.206	-0.978
	(2.151)	(1.803)	(0.944)	(0.147)	(-0.208)	(-1.161)	(-1.357)	(-1.527)	(-1.287)	(-1.011)
Loan size	0.465***	0.526***	0.558***	0.590***	0.602***	0.630***	0.624***	0.627***	0.615***	0.581***
	(4.349)	(4.551)	(4.785)	(4.774)	(5.115)	(4.914)	(4.857)	(4.628)	(4.322)	(3.927)
Deposits	-0.196	-0.303*	-0.338**	-0.343**	-0.504***	-0.578**	-0.579**	-0.511**	-0.488**	-0.474**
	(-1.340)	(-1.915)	(-2.354)	(-2.396)	(-2.782)	(-2.566)	(-2.635)	(-2.386)	(-2.644)	(-2.419)
ROA	-0.053	-0.071**	-0.092***	-0.122***	-0.047**	0.012*	0.020***	0.034***	0.022***	0.021***
	(-1.565)	(-2.122)	(-5.690)	(-4.391)	(-2.360)	(1.866)	(4.499)	(5.431)	(4.049)	(3.571)
MTB	-0.040***	-0.032**	-0.025	-0.012	-0.024*	-0.031**	-0.022	-0.010	0.011	0.021
	(-2.762)	(-2.224)	(-1.488)	(-0.781)	(-1.917)	(-2.165)	(-1.408)	(-0.639)	(0.698)	(1.203)
Reputation	-0.112	-0.310	-0.683	-0.875*	-1.197**	-1.461**	-1.486**	-1.113*	-0.546	0.154
	(-0.205)	(-0.576)	(-1.392)	(-1.752)	(-2.067)	(-2.344)	(-2.254)	(-1.776)	(-0.962)	(0.294)
Specialization in syndicated loan	0.905	0.791	-0.569	-1.554	-1.557	-3.635	-5.793	-7.814	-9.031	-8.628
	(0.389)	(0.365)	(-0.278)	(-0.555)	(-0.569)	(-1.013)	(-1.374)	(-1.531)	(-1.553)	(-1.264)
Specialization in industry	0.045	0.021	-0.013	-0.024	-0.029	-0.058	-0.065	-0.060	-0.040	-0.022
	(0.992)	(0.489)	(-0.389)	(-0.669)	(-0.736)	(-1.143)	(-1.079)	(-0.929)	(-0.636)	(-0.354)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.573	0.603	0.634	0.583	0.567	0.559	0.561	0.572	0.571

Table OA.1.6: **Bank Anomalous Lending and Bank Loan Loss Provisions Controlling for Loan Pricing**

Table OA.1.6 presents the baseline h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, same as in Table 4, additionally controlling for the all-in-drawn spread (in percentage points) weighted by dollar loan amount of syndicate loans the bank participates in over the past 12 months. In all specifications, bank-level A-score and control variables are measured at time t and the dependent variable is the bank's loan loss provisions measured at time $t + h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	0.000	0.002	0.007**	0.010***	0.012***	0.014***	0.013***	0.012**	0.009*	0.004
	(0.119)	(0.500)	(2.455)	(3.311)	(3.936)	(4.024)	(3.072)	(2.518)	(1.853)	(0.708)
Size	0.033***	0.030***	0.027***	0.025***	0.019*	0.016	0.017	0.018	0.018*	0.012
	(3.451)	(3.214)	(3.125)	(2.867)	(1.749)	(1.202)	(1.321)	(1.483)	(1.700)	(1.075)
Tier 1 capital	1.792**	1.079	0.623	0.071	-0.243	-1.124	-1.238	-1.315	-1.099	-0.872
	(2.565)	(1.641)	(0.805)	(0.091)	(-0.305)	(-1.497)	(-1.529)	(-1.565)	(-1.164)	(-0.918)
Loan size	0.463***	0.516***	0.548***	0.590***	0.591***	0.603***	0.569***	0.549***	0.515***	0.479***
	(4.648)	(4.900)	(5.392)	(5.657)	(5.728)	(5.633)	(5.263)	(5.157)	(4.587)	(4.234)
Deposits	-0.151	-0.245	-0.307**	-0.333**	-0.469***	-0.508**	-0.492**	-0.432**	-0.428**	-0.432**
	(-1.160)	(-1.677)	(-2.200)	(-2.225)	(-2.736)	(-2.488)	(-2.597)	(-2.322)	(-2.548)	(-2.222)
ROA	-0.052	-0.070**	-0.090***	-0.122***	-0.046***	0.013***	0.023***	0.037***	0.029***	0.027***
	(-1.481)	(-2.151)	(-4.595)	(-4.574)	(-3.096)	(2.870)	(5.445)	(7.183)	(4.668)	(4.739)
MTB	-0.037***	-0.032***	-0.026*	-0.012	-0.024*	-0.030**	-0.020	-0.009	0.012	0.022
	(-4.309)	(-3.010)	(-1.939)	(-0.822)	(-1.770)	(-2.173)	(-1.321)	(-0.585)	(0.699)	(1.183)
Reputation	-0.224	-0.430	-0.798	-0.941*	-1.292**	-1.617***	-1.618**	-1.325**	-0.850	-0.102
	(-0.402)	(-0.799)	(-1.609)	(-1.896)	(-2.559)	(-2.844)	(-2.569)	(-2.020)	(-1.400)	(-0.186)
Specialization in syndicated loan	1.009	1.003	-0.266	-1.316	-1.289	-3.203	-5.289	-6.947	-8.041	-7.518
	(0.420)	(0.467)	(-0.139)	(-0.486)	(-0.491)	(-0.973)	(-1.409)	(-1.510)	(-1.554)	(-1.207)
Specialization in industry	0.030	0.006	-0.012	-0.024	-0.031	-0.075	-0.083	-0.082	-0.061	-0.055
	(0.630)	(0.136)	(-0.350)	(-0.606)	(-0.741)	(-1.458)	(-1.426)	(-1.405)	(-1.112)	(-1.022)
Loan spread	0.013	0.019*	0.013	0.006	0.012	0.029*	0.041**	0.053**	0.062**	0.066**
	(1.133)	(1.690)	(1.387)	(0.672)	(1.076)	(1.951)	(2.315)	(2.681)	(2.648)	(2.423)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,479	2,372	2,299	2,225	2,152	2,081	2,013	1,950	1,888	1,817
Adjusted R^2	0.556	0.576	0.609	0.633	0.585	0.577	0.576	0.572	0.575	0.573

Table OA.1.7: **Predicting Bank Loan Loss Provisions: Interactions of Categorical A-scores**

Table OA.1.7 presents the baseline h-quarter-ahead prediction results of bank-level quarterly A-scores estimated using the four categories of DealScan loan attributes and their interactions. In all specifications, the bank-level A-scores and control variables are measured at time t. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score (loan covenant): A1	0.052**	0.053*	0.040**	0.044***	0.042**	0.047**	0.045**	0.039*	0.038*	0.047*
	(2.472)	(1.999)	(2.656)	(2.761)	(2.262)	(2.562)	(2.076)	(1.923)	(1.969)	(1.869)
Rank A-score (loan pricing): A2	0.033***	0.038***	0.024***	0.034***	0.019	0.038*	0.046	0.029	-0.002	-0.014
	(2.734)	(2.933)	(3.048)	(2.811)	(0.868)	(1.766)	(1.611)	(1.088)	(-0.088)	(-0.738)
Rank A-score (loan nonpricing): A3	0.061**	0.081**	0.031**	0.027	0.009	0.040	0.062*	0.040	0.038	-0.004
	(2.038)	(2.546)	(2.035)	(1.359)	(0.338)	(1.533)	(1.789)	(1.647)	(1.044)	(-0.196)
Rank A-score (borrower): A4	0.032**	0.030**	0.011	0.017	0.021	0.033**	0.047*	0.027	0.012	0.007
	(2.110)	(2.440)	(1.130)	(1.511)	(1.174)	(2.131)	(1.741)	(0.992)	(0.595)	(0.433)
A1 × A2	-0.003	-0.004	-0.004**	-0.007***	-0.005**	-0.006**	-0.006	-0.004	-0.001	-0.002
	(-1.361)	(-0.980)	(-2.210)	(-3.276)	(-2.061)	(-2.124)	(-1.516)	(-1.392)	(-0.163)	(-0.670)
A1 × A3	-0.011**	-0.013***	-0.007**	-0.006*	-0.004	-0.007*	-0.009*	-0.007**	-0.009*	-0.008
	(-2.629)	(-2.756)	(-2.215)	(-1.755)	(-1.092)	(-1.985)	(-1.959)	(-2.284)	(-1.770)	(-1.627)
A2 × A3	-0.008**	-0.011**	-0.005*	-0.005	-0.000	-0.007	-0.010*	-0.006	-0.003	0.004
	(-2.316)	(-2.464)	(-1.922)	(-1.463)	(-0.043)	(-1.274)	(-1.924)	(-1.375)	(-0.488)	(1.103)
A1 × A4	-0.007**	-0.007*	-0.004	-0.004	-0.006*	-0.007**	-0.008*	-0.005	-0.002	-0.004
	(-2.153)	(-1.907)	(-1.569)	(-1.533)	(-1.726)	(-2.459)	(-2.007)	(-1.411)	(-0.688)	(-0.972)
A2 × A4	-0.004*	-0.004**	-0.001	-0.003	-0.002	-0.006*	-0.010*	-0.004	0.000	0.001
	(-1.738)	(-2.403)	(-0.716)	(-1.261)	(-0.607)	(-1.795)	(-1.760)	(-0.917)	(0.109)	(0.213)
A3 × A4	-0.009	-0.012*	-0.004	-0.005	-0.004	-0.010**	-0.015***	-0.008*	-0.007	0.000
	(-1.642)	(-2.007)	(-1.223)	(-1.293)	(-0.851)	(-2.506)	(-2.766)	(-1.683)	(-1.638)	(0.045)
A1 × A2 × A3	0.001*	0.001*	0.001*	0.001**	0.001	0.001**	0.002*	0.001*	0.001	0.000
	(1.760)	(1.793)	(1.876)	(2.305)	(1.245)	(2.160)	(1.991)	(1.959)	(0.869)	(0.748)
A1 × A2 × A4	0.000	0.001	0.000	0.001*	0.001	0.001**	0.002*	0.001	-0.000	-0.000
	(1.383)	(1.090)	(1.143)	(1.879)	(1.552)	(2.195)	(1.697)	(1.136)	(-0.527)	(-0.080)
A1 × A3 × A4	0.002**	0.002***	0.001*	0.001	0.001	0.002***	0.002***	0.001**	0.001	0.001
	(2.432)	(2.772)	(1.837)	(1.569)	(1.406)	(2.995)	(2.997)	(2.156)	(1.553)	(0.529)
A2 × A3 × A4	0.001	0.002	0.000	0.000	0.000	0.002*	0.003**	0.001	0.000	-0.001
	(1.488)	(1.670)	(0.695)	(0.820)	(0.266)	(1.854)	(2.488)	(1.258)	(0.628)	(-0.608)
A1 × A2 × A3 × A4	-0.000	-0.000*	-0.000	-0.000	-0.000	-0.000**	-0.000**	-0.000*	-0.000	0.000
	(-1.658)	(-1.878)	(-1.140)	(-1.489)	(-1.286)	(-2.678)	(-2.527)	(-1.715)	(-0.230)	(0.310)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.566	0.582	0.605	0.638	0.589	0.571	0.562	0.562	0.576	0.577

Table OA.1.8: **Bank Anomalous Lending and Nonperforming Loans**

Table OA.1.8 presents the h-quarter-ahead prediction results of bank-level A-score for future actual bank nonperforming loans by category. In all specifications, bank-level A-score and control variables are measured at time t, and the dependent variable is the bank's loan loss provisions measured at time $t+h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Nonperforming loans to firms										
Rank A-score	0.012	0.012	0.017*	0.021**	0.027**	0.029***	0.024**	0.023**	0.022*	0.015
	(1.639)	(1.426)	(1.829)	(2.078)	(2.580)	(2.951)	(2.353)	(2.157)	(2.006)	(1.170)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,430	2,356	2,326	2,271	2,193	2,123	2,047	1,999	1,933	1,862
Adjusted R^2	0.594	0.607	0.599	0.594	0.595	0.599	0.590	0.587	0.581	0.570
Nonperforming loans to individuals										
Rank A-score	0.001	0.001	0.000	0.001	0.001	0.002	0.002	0.002	0.002	0.003
	(0.963)	(0.769)	(0.364)	(0.581)	(0.756)	(0.996)	(1.006)	(1.006)	(1.053)	(1.124)
Controls and Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.240	0.237	0.233	0.227	0.226	0.229	0.230	0.233	0.232	0.231

Table OA.1.9: **Interest Income, Anomalous Lending and Loan Loss Provisions**

Table OA.1.9 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, as moderated by the bank's interest-to-noninterest income ratio. We create an indicator High interest-to-noninterest income that takes the value of 1 if the bank's interest-to-noninterest income ratio is above the year-quarter median and 0 otherwise. In all specifications, we include bank-level A-score rank and its interaction with the High interest-to-noninterest income indicator, as well as the same set of control variables as in Table 4 measured at time t. The dependent variable is the bank's loan loss provisions measured at time $t + h$. For simplicity, we do not report the coefficient estimates of the control variables. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h=2	h=4	h=6	h=8	h=10	h=12	h=14	h=16	h=18	h=20
Rank A-score	-0.001	-0.004	-0.001	-0.000	0.002	0.001	0.002	0.000	-0.003	-0.008
	(-0.254)	(-0.696)	(-0.179)	(-0.094)	(0.442)	(0.221)	(0.471)	(0.032)	(-0.600)	(-1.343)
Rank A-score × High interest-to-noninterest income	0.003	0.009	0.013*	0.017**	0.017**	0.024***	0.020***	0.022***	0.020**	0.020**
	(0.383)	(1.181)	(1.928)	(2.511)	(2.590)	(4.059)	(3.506)	(3.671)	(2.472)	(2.045)
High interest-to-noninterest income	0.013	-0.003	-0.028	-0.050	-0.046	-0.068	-0.044	-0.037	-0.012	0.002
	(0.310)	(-0.062)	(-0.666)	(-1.240)	(-0.974)	(-1.527)	(-0.951)	(-0.854)	(-0.244)	(0.037)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.575	0.605	0.637	0.586	0.573	0.563	0.568	0.579	0.580

Table OA.1.10: **Global Financial Crisis, Anomalous Lending and Loan Loss Provisions**

Table OA.1.10 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, as moderated by the Global Financial Crisis (GFC) period from 2007Q1 to 2009Q4. In all specifications, we include bank-level A-score rank and its interaction with the GFC indicator, as well as the same set of control variables as in Table 4 measured at time t. The dependent variable is the bank's loan loss provisions measured at time $t + h$. For simplicity, we do not report the coefficient estimates of the control variables. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	0.001	0.000	0.003	0.005**	0.006**	0.010***	0.010**	0.012*	0.008	0.004
	(0.295)	(0.163)	(1.460)	(2.059)	(2.191)	(3.478)	(2.393)	(1.832)	(1.258)	(0.520)
Rank A-score × GFC	0.001	0.014	0.028	0.038***	0.043***	0.031*	0.022*	0.004	0.002	0.000
	(0.055)	(0.552)	(1.490)	(2.872)	(3.441)	(1.757)	(1.933)	(0.484)	(0.346)	(0.047)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.574	0.607	0.641	0.593	0.572	0.561	0.561	0.572	0.571

Table OA.1.11: **Sample Period, Anomalous Lending and Loan Loss Provisions**

Table OA.1.11 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions. Specifically, we split the sample period into two equal-length sub-periods and create the PostMid indicator that takes the value of 1 for observations in the more recent sub-sample and 0 otherwise. In all specifications, we include bank-level A-score rank and its interaction with the PostMid indicator, as well as the same set of control variables as in Table 4 measured at time t. The dependent variable is the bank's loan loss provisions measured at time $t + h$. For simplicity, we do not report the coefficient estimates of the control variables. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	0.001	0.002	0.006**	0.009***	0.012***	0.013***	0.013**	0.012*	0.009	0.004
	(0.411)	(0.668)	(2.466)	(3.476)	(3.855)	(3.352)	(2.429)	(1.979)	(1.486)	(0.599)
Rank A-score × PostMid	-0.002	0.008	0.006	0.014	0.001	0.022	0.009	0.005	-0.004	-0.001
	(-0.067)	(0.241)	(0.575)	(1.262)	(0.086)	(1.296)	(0.604)	(0.504)	(-0.593)	(-0.155)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.573	0.603	0.634	0.583	0.568	0.559	0.561	0.572	0.571

Table OA.1.12: **Overview of Stress Test Outcomes, 2009-2019**

Table OA.1.12 provides an overview of the participating bank holding companies and outcomes of the stress tests across different rounds. 1 indicates that the bank's capital plan received an objection or conditional non-objection and 0 indicates otherwise. Shaded cells indicate banks that are not eligible to participate in a focal round of stress test. Note that Deutsche Bank launched its US intermediate holding company, DB USA Corporation, on July 1, 2016, under which most of its US-based operations were consolidated. MUFG Americas Holdings Corporation was formerly UnionBanCal Corporation before 2014. In February 2019, the Federal Reserve announced that certain banks with total consolidated assets between $100 billion and $250 billion would not be subject to the company-run and supervisory stress testing requirements nor the requirement to submit a capital plan during the 2019 cycle.

Bank	RSSD ID	SCAP 2009	CCAR 2012	CCAR 2013	CCAR 2014	CCAR 2015	CCAR 2016	CCAR 2017	CCAR 2018	CCAR 2019
Ally Financial Inc.	1562859	1	1	1	0	0	0	0	0	
American Express Company	1275216	0	0	0	0	0	0	0	0	
Bank of America Corporation	1073757	1	0	0	0	1	0	0	0	0
BancWest Corporation	1025608						0	0		
Barclays US LLC	5006575								0	0
BB&T Corporation	1074156	0	0	1	0	0	0	0	0	
BBVA Compass Bancshares, Inc.	1078529				0	0	0	0	0	
BMO Financial Corp.	1245415				0	0	0	0	0	
BNP Paribas USA, Inc.	1575569								0	
Capital One Financial Corporation	2277860	0	0	0	0	0	0	1	0	0
CIT Group Inc.	1036967							0		
Citigroup Inc.	1951350	1	1	0	1	0	0	0	0	0
Credit Suisse Holdings (USA), Inc.	1574834									1
Comerica Incorporated	1199844				0	0	0	0	0	
Deutsche Bank Trust Corporation	1032473					1	1	0		
DB USA Corporation	2816906								1	0
Discover Financial Services	3846375				0	0	0	0	0	
Fifth Third Bancorp	1070345	1	0	0	0	0	0	0	0	
HSBC North America Holdings Inc.	3232316				1	0	0	0	0	0
Huntington Bancshares Incorporated	1068191				0	0	0	0	0	
JPMorgan Chase & Co.	1039502	0	0	1	0	0	0	0	0	0
Keycorp	1068025	1	0	0	0	0	0	0	0	
M&T Bank Corporation	1037003				0	0	0	0	0	
MetLife, Inc.	2945824	0	1							
Morgan Stanley	2162966	1	0	0	0	0	1	0	1	0
MUFG Americas Holdings Corporation	1378434				0	0	0	0	0	
Northern Trust Corporation	1199611				0	0	0	0	0	0
RBC USA Holdco Corporation	3226762								0	
RBS Citizens Financial Group, Inc.	1132449				1	0	0	0	0	
Regions Financial Corporation	3242838	1	0	0	0	0	0	0	0	
Santander Holdings USA, Inc.	3981856				1	1	1	0	0	
State Street Corporation	1111435	0	0	0	0	0	0	0	1	0
SunTrust Banks, Inc.	1131787	1	1	0	0	0	0	0	0	
TD Group US Holdings LLC	3606542						0	0	0	0
The Bank of New York Mellon Corporation	3587146	0	0	0	0	0	0	0	0	0
The Goldman Sachs Group, Inc.	2380443	0	0	1	0	0	0	0	1	0
The PNC Financial Services Group, Inc.	1069778	1	0	0	0	0	0	0	0	0
U.S. Bancorp	1119794	0	0	0	0	0	0	0	0	0
UBS Americas Holding LLC	4846998									0
Wells Fargo & Company	1120754	1	0	0	0	0	0	0	0	0
Zions Bancorporation	1027004				1	0	0	0	0	

OA.2. Estimating A-score

The A-score measure of bank anomalous lending is computed by aggregating the loan-level anomaly scores estimated by applying the Isolation Forest algorithm on the syndicated loan data from DealScan. This Appendix details our estimation process.

To apply Isolation Forest algorithm, we construct a training dataset at each quarter's end of all syndicated loans originated in the past 12 months. Starting with an inspection of all 203 variables in the new consolidated DealScan dataset which contains lender-loan-level observations from 1990 to 2021, we manually screen and select 152 variables. Table OA.2.3 tabulates all DealScan variables and whether and why they are included in the estimation. We detail the individual reason for excluded variables. The most common reasons for excluding a variable include, for example, 1) the variable is one-to-one mapped to another variable (e.g., firm name and ID), 2) the variable has no variation (e.g., lender's country being U.S.), 3) the variable provides no additional information (e.g., some are only concatenation of other variables), and 4) the variable contains *almost all* missing values.[25]

Next, we convert a subset of the 203 variables from numeric or textual to indicators. For example, we convert 'Additional Borrowers', which lists the additional borrowers' names, if any, to an indicator 'Has Additional Borrowers' to capture the existence of extra borrowers; 'Purpose Remark' is converted to an indicator 'Has Purpose Remark' to reflect whether there is additional comment; 'CUSIP' is converted to an indicator 'Has CUSIP' to distinguish loans with and without CUSIP assigned, and so on. We also convert loan covenants to indicators. For example, 'Max Leverage Ratio' is converted into an indicator 'Has Max

[25]We note that even variables with a significant proportion of missing values are still informative and kept in our training sample, because having non-missing data whilst others do not is 'abnormal' as well. Variables are excluded if we have only a couple hundreds non-missing values out of the 584,241 total observations.

Leverage Ratio Covenant'. Moreover, we transform certain textual variables into categorical to better extract useful information. For example, the 'Repayment Schedule' variable describes the repayment schedule such as '3 Annually USD 333.333m 8/3/2011', '1 Bullet USD 6m 12/1/2020', etc. We identify all repayment frequencies and convert the variable into 'monthly', 'quarterly', 'annually', 'variable', 'bullet', 'biannually', and 'other'. For date variables, we use the 'Tranche Active Date' as the benchmark and convert all other dates to days relative to the benchmark, thereby making comparable the time differences. We also manually check and correct 91 law firm names. For example, Clifford Chance (UK), Clifford Chance Hong Kong, Clifford Chance US LLP and Clifford Chance Pünder are all converted to Clifford Chance LLP. Typos in DealScan such as Stephenson Harwood incorrectly entered as Stephenson Hardwood are also corrected. We document in the notes of Table OA.2.3 the conversion and/or correction for each variable.

A special scenario occurs when a variable can have multiple categorical values at the same time. For example, the 'Lead Arranger' variable may contain the information of the lead arranger banks and their shares, such as 'Banco Bilbao Vizcaya Argentaria SA [BBVA] 5.73%, HypoVereinsbank 5.73%, Royal Bank of Scotland Plc [RBS]', 'Barclays Bank Plc, Clydesdale Bank Plc, HSBC Bank Plc, Rabobank', and 'Pioneer Bank (Kentucky) 100.00%'. To capture the lenders involved in the syndicate, we take advantage of binary encoding to convert 'Lead Arranger' into a numeric variable. Specifically, we sort all lenders in the sample alphabetically and assign the ith lender a code of 2^{i-1}, which in binary form has a value of 1 at the ith digit and 0 at all other places. Table OA.2.1 presents an example encoding assuming 5 lenders in the sample.

Table OA.2.1: **Example Encoding of Each Lender**

ID	Lender	Decimal Representation	Binary Representation
1	Bank A	$2^0 = 1$	00000001
2	Bank B	$2^1 = 2$	00000010
3	Bank C	$2^2 = 4$	00000100
4	Bank D	$2^3 = 8$	00001000
5	Bank E	$2^4 = 16$	00010000

Now, suppose a loan syndicate has bank A and B as lead arrangers, we encode 'Lead Arranger' as $2^0 + 2^1$ based on Table OA.2.1, whose binary representation will have '1' at the first and second digits, thereby indicating the presence of the two lenders. Table OA.2.2 shows some more examples of how this encoding process works, where $\|$ is the binary OR operator.

Table OA.2.2: **Example Encoding of Lead Arranger**

Loan	Lead Arranger	Binary Encoding	Decimal Encoding
1	1. Bank A	00000001	$2^0 = 1$
2	1. Bank A 2. Bank B	00000001 $\|$ 00000010 = 00000011	$2^0 + 2^1 = 3$
3	1. Bank A 2. Bank B 3. Bank C	00000001 $\|$ 00000010 $\|$ 00000100 = 00000111	$2^0 + 2^1 + 2^2 = 7$
4	2. Bank B 4. Bank D	00000010 $\|$ 00001000 = 00001010	$2^2 + 2^4 = 20$

Note that this encoding is bi-directional as we know the lenders from the encoding. We can therefore use a single numeric value to represent all lead arrangers and enable comparison in the Isolation Forest algorithm. The same binary encoding method is applied on all

other multi-categorical variables, such as 'Bookrunner', 'Top Tier Arranger', 'Lead Left', 'Arranger', 'Co Arranger', 'Agent', 'Lead Manager', 'All Lenders', 'Organization Type', 'Market Segment', etc.

At the end of each quarter, we construct a training dataset consisted of all loans originated in the past 36 months. We then apply the Isolation Forest algorithm to compute the anomaly score for each loan as of the quarter's end. The bank-level A-score is then computed following the description in Section 4.

Table OA.2.3: **DealScan Variables Used in Estimation**

Variable	Included	Reason	Note
Category: Borrower			
Borrower_Name	No	1-to-1 mapped to Borrower_Id	
Borrower_Id	Yes		
Additional_Borrowers	Yes		Converted to an indicator 'Has_Additional_Borrowers'.
Ticker	No	1-to-1 mapped to Borrower_Id	
Perm_ID	No	1-to-1 mapped to Borrower_Id	
Legal_Entity_ID_LEI	No	1-to-1 mapped to Borrower_Id	
City	Yes		
State_Province	Yes		
Zip	Yes		
Country	Yes		
Region	Yes		
Sales_Size	Yes		
Broad_Industry_Group	Yes		
Major_Industry_Group	Yes		
SIC_Code	Yes		
NAIC	Yes		
Borrower_Type	Yes		
Parent	Yes		
Parent_Ticker	No	1-to-1 mapped to borrower's Parent	
Sponsor	Yes		
Guarantor	Yes		
Target	No		
Organization_Type	Yes		Can be 'Company', 'Listed Company', 'SPV', 'Government-owned', etc.
Senior_Debt_to_EBITDA	Yes		
Total_Debt_to_EBITDA	Yes		
Company_Url	Yes	1-to-1 mapped to borrower. But the domain info can be used. '.com', '.net', '.jp', '.ru', etc.	Top-level domain (TLD) names used.
Sales_Size_at_Close	Yes		The financial amount by which the company's sales (revenue) is measured as of the closing date of the agreement
Accounts_Receivable	Yes		
Oil_Gas_Reserves	No	Almost all missing	
Acc_Rec_Domestic	No	Almost all missing	
Acc_Rec_Foreign	No	Almost all missing	
Inventory	Yes		
Inv_Raw_Material	No	Almost all missing	
Inv_Finished_Goods	No	Almost all missing	
Inv_Work_in_Progress	No	Almost all missing	
Cash_Cash_Equivalents	No	Almost all missing	
Property_Plant_Equipment	No	Almost all missing	
Marketable_Securities	No	Almost all missing	
Eligible_Property_Value	No	Almost all missing	

Table OA.2.3: **Continued**

Variable	Included	Reason	Note
Category: Loan-nonpricing			
Lead_Arranger	Yes		
Number_of_Lead_Arrangers	Yes		
Bookrunner	Yes		
Number_of_Bookrunners	Yes		
Top_Tier_Arranger	Yes		
Number_of_Top_Tier_Arrangers	Yes		
Lead_Left	Yes		
Number_of_Lead_Left	Yes		
Arranger	Yes		
Number_of_Arrangers	Yes		
Co_Arranger	Yes		
Number_of_Co_Arrangers	Yes		
Agent	Yes		
Number_of_Agents	Yes		
Lead_Manager	Yes		
Number_of_Lead_Managers	Yes		
All_Lenders	Yes		
Number_of_Lenders	Yes		
Lender_Region	Yes		
Lender_Parent_Region	Yes		
Lender_Institution_Type	Yes		
LPC_Deal_ID	No	1-to-1 mapped to Deal_PermID	
Deal_PermID	Yes	Inclusion of this identifier can enable comparison within a deal	
Deal_Amount	No	Use Deal_Amount_Converted instead	
Deal_Amount_Converted	Yes		
Deal_Active_Date	Yes		
Deal_Currency	Yes		
Deal_Input_Date	No	Date of the record entering DealScan	
Deal_Purpose	Yes		
Project_Finance	No	Missing a lot. It lists the type of the project.	We have 'Project Finance' in the 'Primary_Purpose' as a category so don't need to make a new variable, say 'Is_Project_Finance'.
Phase	Yes		
Deal_Active	Yes	Correlated with Phase, loan duration and active date. But included.	
Deal_Refinancing	Yes		
Purpose_Remark	Yes		Convert to an indicator 'Has_Purpose_Remark'
Deal_Remark	Yes		Convert to an indicator 'Has_Deal_Remark'
Deal_Amended	Yes		
LPC_Tranche_ID	No	1-to-1 mapped to Tranche_PermID	
Tranche_PermID	Yes		
Tranche_Type	Yes		
Tranche_Active_Date	Yes		
Tranche_Maturity_Date	No	Use Tenor_Maturity instead	
Tranche_Amount	No	Use Tranche_Amount_Converted instead	
League_Table_Amount	No	Use League_Table_Amount_Converted instead	
League_Table_Amount_Converted	Yes		
_100_percent_Vote	Yes		
Tranche_O_A	No	Only 'Origination' used	
Tranche_Amount_Converted	Yes		
Tranche_Currency	Yes		
Tranche_CUSIP	Yes		Convert to an indicator 'Has_CUSIP'
Market_Segment	Yes		
Market_Of_Syndication	Yes		
Country_Of_Syndication	Yes		
League_Table_Credit	Yes		
League_Table_Tranche_Date	Yes		
Seniority_Type	Yes		
Secured	Yes		
Repayment_Type	Yes		
Repayment_Schedule	Yes		Convert to 'Quarterly', 'Monthly', 'Bullet', etc. If 'variable payment' is present, it is labelled as 'Variable'. In 14 cases the repayment schedule is 'every two years' and converted to 'Biannually'. The rests are labelled as 'Other'.

Table OA.2.3: **Continued**

Variable	Included	Reason	Note
Category: Loan-nonpricing (continued)			
Distribution_Method	Yes		
Tranche_Remark	Yes		Convert to an indicator 'Has_Tranche_Remark'
Primary_Purpose	Yes		
Secondary_Purpose	Yes		
Tertiary_Purpose	Yes		Convert to an indicator 'Has_Tertiary_Purpose'
Tranche_Refinancing	Yes		
Sponsored	Yes		
Project_Finance_Sponsor	Yes		Convert to an indicator 'Has_Project_Finance_Sponsor'
Tranche_Active	Yes		
Tenor_Maturity	Yes		
Closed_Date	Yes		Convert to days relative to 'Tranche_Active_Date'
Completion_Date	Yes		Convert to days relative to 'Tranche_Active_Date'
Mandated_Date	Yes		Convert to days relative to 'Tranche_Active_Date'
Launch_Date	Yes		Convert to days relative to 'Tranche_Active_Date'
Average_Life	No	Use Tenor_Maturity instead	
Multi_Currency_Tranche	Yes		
Borrower_Consent	Yes		
Agent_Consent	Yes		
Assignment_Minimum	Yes		
Assignment_Fee	Yes		
New_Money	No	Use New_Money_Converted instead	
New_Money_Converted	Yes		
Amend_Extend_Flag	Yes		
Tranche_Amended	Yes		
Currency_Onshore_Offshore	Yes		
Pro_Rata	Yes		
LIN	Yes		Convert to an indicator 'Has_LIN'
Letter_Of_Credit	Yes		
Swingline	Yes		Dollar amount for swingline sublimit within the facility
Multi_Currency	Yes		Dollar amount for multi-currency drawing sublimit within the facility.
Bid_Option	Yes		Flag indicating if the facility has a Competitive Bid option/rate option
Bankers_Acceptance	Yes		A flag indicating if the facility has a Banker's Acceptance option/rate option.
Foreign_Exchange	No	Almost all missing	Dollar amount for foreign exchange drawing sublimit within the facility.
Law_Firm_Name	No	A concatenation of vars from Law_Firm_Lender_Primary to Law_Firm_Borrower_Other	
Law_Firm_Lender_Primary	Yes		Manually checked and corrected 91 law firm names. For example, Clifford Chance (UK), Clifford Chance Hong Kong, Clifford Chance US LLP and Clifford Chance Punder are all converted to Clifford Chance LLP. Typos in DealScan such as Stephenson Harwood incorrectly entered as Stephenson Hardwood are also corrected.
Law_Firm_Lender_Other	No	Focus on just the law firm of the primary lender	
Law_Firm_Borrower_Primary	Yes		Same manual check as above.
Law_Firm_Borrower_Other	No	Focus on just the law firm of the primary borrower	

Table OA.2.3: **Continued**

Variable	Included	Reason	Note
Category: Loan-covenant			
Collateral_Security_Type	Yes		
Covenants	Yes		Convert to indicators
Max_Leverage_Ratio	Yes		Convert to indicators
Max_Debt_to_Cash_Flow	Yes		Convert to indicators
Max_Sr_Debt_to_Cash_Flow	Yes		Convert to indicators
Tangible_Net_Worth	Yes		Convert to indicators
Net_Worth	Yes		Convert to indicators
Min_Fixed_Charge_Coverage_Ratio	Yes		Convert to indicators
Min_Debt_Service_Coverage_Ratio	Yes		Convert to indicators
Min_Interest_Coverage_Ratio	Yes		Convert to indicators
Min_Cash_Interest_Coverage_Ratio	Yes		Convert to indicators
Max_Debt_to_Tangible_Net_Worth	Yes		Convert to indicators
Max_Debt_to_Equity_Ratio	Yes		Convert to indicators
Min_Current_Ratio	Yes		Convert to indicators
Max_Loan_to_Value_Ratio	Yes		Convert to indicators
All_Covenants_Financial	Yes		Convert to indicators
Covenant_Comment	Yes		Convert to an indicator 'Has_Covenant_Comment'
Excess_CF_Sweep	Yes		The percentage amount of net proceeds a borrower receives from excess CF that must be used to reduce any loan balance outstanding.
Asset_Sales_Sweep	Yes		The percentage amount of net proceeds a company receives from an asset sale that must be used to pay down any outstanding loan balance.
Material_Restriction	Yes		
Debt_Issue_Sweep	Yes		The percentage amount of net proceeds a company receives from the issuance of debt that must be used to pay down any outstanding loan balance.
Equity_Issue_Sweep	Yes		The percentage amount of net proceeds a company receives from the issuance of equity that must be used to pay down any outstanding loan balance.
Insurance_Proceeds_Sweep	Yes		The percentage amount of net proceeds a company receives from insurance settlements that must be used to pay down any outstanding loan balance.
Percentage_of_Net_Income	Yes		Percentage of net income the borrower is allowed t use towards dividends.
Collateral_Release	Yes		The percentage amount of lenders required to release the borrower from collateral covenants.
Required_Lenders	Yes		The percentage amount of the deal commitment lenders must be holding in order to approve any modifications to the present deal.
Terms_Changes	Yes		The percentage amount of lenders required to release the borrower from terms
EBITDA_Initial_Amount_USD	Yes		
EBITDA_Final_Amount_USD	Yes		
CAPEX_Initial_USD	Yes		
CAPEX_Final_USD	Yes		
All_Covenants_General	No	A concatenation of variables from Asset_Sale_Sweep to CAPEX_Final_USD	

Table OA.2.3: **Continued**

Variable	Included	Reason	Note
Category: Loan-pricing			
Base_Reference_Rate	Yes		
Margin_bps	Yes		
All_Base_Rate_Spread_Margin	Yes	A concatenation of Base_Reference_Rate and Margin_bps, plus other spreads, if any.	Convert to an indicator 'Has_Multiple_Spread'
Base_Rate_Margin_bps	No	A concatenation of Base_Reference_Rate and Margin_bps	
Floor_bps	Yes		Convert to an indicator 'Has_Floor_bps'
Original_Issue_Discount	Yes		
Call_Protection	Yes		
Call_Protection_Text	No	Details when Call_Protection=TRUE	
All_In_Spread_Drawn_bps	Yes		
All_in_Spread_Undrawn_bps	Yes		
Annual_Fee_bps	Yes		
Commitment_Fee_bps	Yes		
Letter_Of_Credit_Fee_bps	Yes		
Upfront_Fee_bps	Yes		
All_in_Fee_Asia_only	No		
Base_Rate_Comment	No	Duplicate Base_Reference_Rate	
Repayment_Comment	No		
Performance_Pricing	No	Use the indicator 'Performance_Pricing_Grid' instead	
Performance_Pricing_Grid	Yes		
Performance_Pricing_Remark	Yes		Convert to an indicator 'Has_Performance_Pricing_Remark'
Cancellation_Fee_bps	Yes		
Documentary_Issuing_Fee_bps	No	Almost all missing	
Documentary_LC_Fee_bps	No	Almost all missing	
Tiered_Upfront_Fee	Yes		Convert to an indicator 'Has_Tiered_Upfront_Fee'
Utilization_Fee_bps	Yes		
All_Fees	No	Combine all fees together	
Not categorized			
Lender_Parent_Name	No	Not used because the training sample requires that each loan should have only one observation for a certain attribute, while each loan can have multiple lenders	
Lender_Parent_Id	No	Same as above	
Lender_Name	No	Same as above	
Lender_Id	No	Same as above	
Primary_Role	No	Same as above	
Additional_Roles	No	Same as above	
Lender_Commit	No	Same as above	
Lender_Share	No	Same as above	
Lender_Operating_Country	No	Same as above	
Lender_Parent_Operating_Country	No	Same as above	

OA.3. Further results

This Online Appendix provides additional robustness results.

Table OA.3.1 presents the results of the baseline model where the raw bank-level A-score is used instead of the quarterly decile rank of A-score.

Table OA.3.2 presents the results of the baseline model where the A-score is estimated with alternative parameters in the Isolation Forest with a tree size of 512 and a forest size of 5,000.

Table OA.3.1: **Bank Anomalous Lending and Bank Loan Loss Provisions: Raw A-score**

Table OA.3.1 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions. In all specifications, bank-level A-score and control variables are measured at time t, and the dependent variable is the bank's loan loss provisions measured at time $t+h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
A-score	0.292	0.806	2.223***	3.041***	3.532***	4.434***	3.813**	4.529*	3.385	1.512
	(0.299)	(0.911)	(3.723)	(4.657)	(3.775)	(3.825)	(2.260)	(2.008)	(1.570)	(0.775)
Size	0.028***	0.025**	0.024***	0.024***	0.016	0.010	0.010	0.010	0.010	0.006
	(2.779)	(2.512)	(2.749)	(2.799)	(1.445)	(0.653)	(0.603)	(0.623)	(0.665)	(0.372)
Tier 1 capital	1.981***	1.324*	0.711	0.097	-0.179	-0.892	-1.121	-1.326	-1.217	-0.980
	(2.754)	(1.884)	(0.899)	(0.124)	(-0.220)	(-1.154)	(-1.336)	(-1.559)	(-1.265)	(-0.996)
Loan size	0.465***	0.525***	0.558***	0.591***	0.602***	0.628***	0.624***	0.629***	0.616***	0.581***
	(4.739)	(5.121)	(5.822)	(6.128)	(6.286)	(6.418)	(6.331)	(6.161)	(5.645)	(4.984)
Deposits	-0.196	-0.304**	-0.339**	-0.344**	-0.504***	-0.577***	-0.579***	-0.509**	-0.481***	-0.469**
	(-1.419)	(-2.040)	(-2.536)	(-2.446)	(-2.997)	(-2.689)	(-2.753)	(-2.482)	(-2.768)	(-2.491)
ROA	-0.053	-0.071**	-0.092***	-0.122***	-0.047***	0.011**	0.019***	0.034***	0.022***	0.021***
	(-1.529)	(-2.201)	(-4.670)	(-4.568)	(-3.030)	(2.149)	(4.468)	(5.294)	(4.487)	(3.779)
MTB	-0.040***	-0.032***	-0.025*	-0.011	-0.023*	-0.030**	-0.021	-0.009	0.012	0.021
	(-4.460)	(-3.152)	(-1.957)	(-0.817)	(-1.757)	(-2.227)	(-1.401)	(-0.615)	(0.696)	(1.184)
Reputation	-0.107	-0.322	-0.646	-0.808	-1.057**	-1.342**	-1.314**	-1.106*	-0.582	0.128
	(-0.198)	(-0.614)	(-1.343)	(-1.628)	(-2.075)	(-2.521)	(-2.341)	(-1.881)	(-1.020)	(0.230)
Specialization in syndicated loan	0.940	0.855	-0.272	-1.112	-0.967	-2.978	-5.123	-7.357	-8.750	-8.525
	(0.414)	(0.405)	(-0.138)	(-0.417)	(-0.378)	(-0.906)	(-1.348)	(-1.511)	(-1.568)	(-1.268)
Specialization in industry	0.045	0.021	-0.010	-0.020	-0.022	-0.048	-0.056	-0.060	-0.040	-0.023
	(0.994)	(0.488)	(-0.301)	(-0.540)	(-0.535)	(-0.929)	(-0.914)	(-0.906)	(-0.630)	(-0.363)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.573	0.603	0.634	0.582	0.566	0.557	0.562	0.573	0.571

Table OA.3.2: **Bank Anomalous Lending and Bank Loan Loss Provisions: Alternative Parameters**

Table OA.3.2 presents the h-quarter-ahead prediction results of bank-level A-score for future bank loan loss provisions, where the A-score is estimated using alternative parameters in the Isolation Forest with a tree size of 512 and a forest size of 5,000. In all specifications, bank-level A-score and control variables are measured at time t, and the dependent variable is the bank's loan loss provisions measured at time $t + h$. Heteroskedasticity-robust standard errors are clustered at the bank level. Definitions of the variables are provided in Online Appendix OA.1.1. Numbers in parentheses are two-tailed t-statistics. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) h=2	(2) h=4	(3) h=6	(4) h=8	(5) h=10	(6) h=12	(7) h=14	(8) h=16	(9) h=18	(10) h=20
Rank A-score	-0.000	0.002	0.007**	0.010***	0.012***	0.014***	0.014**	0.013**	0.009	0.004
	(-0.042)	(0.464)	(2.582)	(3.291)	(3.285)	(3.334)	(2.514)	(2.012)	(1.481)	(0.583)
Size	0.028***	0.025**	0.024***	0.024***	0.016	0.009	0.010	0.010	0.010	0.005
	(2.763)	(2.517)	(2.802)	(2.869)	(1.467)	(0.627)	(0.610)	(0.629)	(0.636)	(0.357)
Tier 1 capital	1.982***	1.327*	0.715	0.101	-0.173	-0.897	-1.124	-1.309	-1.214	-0.981
	(2.755)	(1.890)	(0.901)	(0.128)	(-0.208)	(-1.134)	(-1.321)	(-1.519)	(-1.255)	(-0.995)
Loan size	0.465***	0.525***	0.557***	0.590***	0.600***	0.628***	0.622***	0.624***	0.613***	0.580***
	(4.742)	(5.112)	(5.731)	(5.983)	(6.094)	(6.208)	(6.089)	(5.935)	(5.503)	(4.944)
Deposits	-0.196	-0.300*	-0.330**	-0.332**	-0.491***	-0.563**	-0.563***	-0.495**	-0.477***	-0.469**
	(-1.400)	(-1.992)	(-2.458)	(-2.377)	(-2.960)	(-2.671)	(-2.749)	(-2.457)	(-2.738)	(-2.472)
ROA	-0.053	-0.071**	-0.092***	-0.123***	-0.048***	0.010**	0.019***	0.033***	0.022***	0.020***
	(-1.533)	(-2.203)	(-4.738)	(-4.644)	(-3.152)	(2.111)	(4.476)	(5.133)	(4.366)	(3.716)
MTB	-0.040***	-0.033***	-0.026*	-0.013	-0.025*	-0.032**	-0.022	-0.011	0.011	0.021
	(-4.484)	(-3.152)	(-2.006)	(-0.899)	(-1.848)	(-2.272)	(-1.463)	(-0.679)	(0.648)	(1.171)
Reputation	-0.048	-0.292	-0.717	-0.913*	-1.144**	-1.390**	-1.466**	-1.108*	-0.537	0.170
	(-0.082)	(-0.521)	(-1.422)	(-1.740)	(-2.144)	(-2.430)	(-2.382)	(-1.792)	(-0.907)	(0.299)
Specialization in syndicated loan	1.020	0.779	-0.793	-1.855	-1.728	-3.837	-6.082	-8.127	-9.234	-8.690
	(0.428)	(0.359)	(-0.391)	(-0.648)	(-0.633)	(-1.095)	(-1.468)	(-1.574)	(-1.604)	(-1.267)
Specialization in industry	0.047	0.020	-0.019	-0.032	-0.034	-0.062	-0.071	-0.068	-0.044	-0.023
	(1.003)	(0.460)	(-0.558)	(-0.875)	(-0.821)	(-1.138)	(-1.097)	(-0.985)	(-0.667)	(-0.350)
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,563	2,443	2,373	2,292	2,215	2,144	2,068	1,999	1,933	1,862
Adjusted R^2	0.557	0.573	0.603	0.634	0.582	0.566	0.558	0.561	0.572	0.571